

Hanna Kukkonen

**TEKOÄLY JA MONIMUTKAINEN
ONGELMANRATKAISU**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2023

TIIVISTELMÄ

Kukkonen, Hanna

Tekoäly ja monimutkainen ongelmanratkaisu

Jyväskylä: Jyväskylän yliopisto, 2023, 86 s

Kognitiotiede, pro gradu -tutkielma

Ohjaaja: Jokinen, Jussi

Tekoälytutkimuksessa on vuosikymmeniä ollut kunnianhimoisena tavoitteena itsenäisesti toimivien älykkäiden kokonaisuuksien luominen. Viime vuosien myönteisestä kehityksestä huolimatta tekoälyn ongelmanratkaisukyvyt ovat edelleen ihmistä heikommät, kun kyseessä on monimutkainen tai jäsentymätön tilanne.

Tässä tutkimuksessa on selvitetty systemaattisen kirjallisuuskatsauksen menetelmällä, mitä keskeisiä ongelmia tekoälyn ongelmanratkaisuun liittyy ja minkälaisia ratkaisuehdotuksia niihin esitetään. Kirjallisuushaun tietokantana toimi scholars.google.fi ja hakusanoina käytettiin termejä *cognitive science solutions for ai complex problem solving*. Tarkennettuina hakukriteereinä olivat tarkka ilmaus *artificial intelligence problem solving* ja julkaisuajankohta 2017–2023. Kirjallisuushaussa löytyi yhteensä 86 artikkelia, joista seulonnan jälkeen tarkasteluun jäi kahdeksan tutkimusta.

Tutkimusaineistosta tunnistetut tekoälyn ongelmanratkaisukyvyn keskeiset haasteet on jaoteltu teorian pohjalta frame-ongelman ja tiedon hankinnan ongelman kategorioihin. Frame-ongelma sisältää näkökulmat ongelman ja tavoite-tilan määrittelyn vaikeudesta sekä kaikkien mahdollisten tilojen määrittelyn haasteen. Tiedon hankinnan ongelma käsittää tiedon mallintamisen ja luotettavuuden näkökulmat. Ratkaisuna em. ongelmiin tutkimusaineistossa esitettiin muun muassa erilaisiin oppimisstrategioihin, kuten siirto-oppiminen, aktiivinen oppiminen ja kokemuspohjainen oppiminen perustuvia malleja, mutta myös hakuvaruuden mallintamista ja mittaamista. Tutkimusaineiston kaikkien artikkelien ratkaisuehdotukset olivat hieman erilaisia, keskeisimpänä yhdistävänä tekijänä oli ihmisen hyödyntäminen tekoälyn tukena joko kouluttajana, aineiston validoijana tai ongelmatilanteen jäsentäjänä. Tutkimuksen johtopäätöksenä esitetään viitekehys, jossa aineistosta nousseet ongelmat on suhteutettu ihmisen ongelmanratkaisun kognitiiviseen prosessiin.

Tutkimus ei ole kaiken kattava katsaus tekoälyn ongelmanratkaisuun ja mahdollisiin ratkaisuihin, mutta se luo suuntaa tulevista kehitysnäkymistä sekä jäsentää tutkimusalaa kognitiotieteellisestä näkökulmasta. Tutkimuksen perusteella vaikuttaa siltä, että yhtä selkeää ratkaisua tekoälyn ongelmanratkaisukyvyn kehittämiseksi ei ole vielä kehitetty eli jatkotutkimukselle on tarvetta. Jatko-tutkimuksena voisi selvittää yleistyvätkö tutkimuksen tulokset kattavammalla aineistolla tai kartoittaa tunnistettujen ratkaisujen mahdollisuuksia ja heikkouksia.

Asiasanat: Älykkyys, tekoäly, monimutkainen ongelmanratkaisu, kognitiotiede, systemaattinen kirjallisuuskatsaus

ABSTRACT

Kukkonen, Hanna

Artificial intelligence and complex problem solving

Jyväskylä: University of Jyväskylä, 2023, 86 pp

Cognitive science, Master's Thesis

Supervisor: Jokinen, Jussi

For decades the ambitious goal of artificial intelligence research has been to create autonomously functioning intelligent entities. Despite the recent years positive development, the problem-solving abilities of artificial intelligence are still inferior to humans in complex or ill-structured situations.

This study examined the key challenges associated with artificial intelligence problem-solving and explored proposed solutions through a systematic literature review. The literature search was conducted on scholars.google.fi using the keywords *cognitive science solutions for AI complex problem-solving*. The refined search criteria included the precise phrase *artificial intelligence problem solving* and the publication range of 2017–2023. A total of 86 articles were found, and after screening, eight studies were included in the review.

The identified key challenges in artificial intelligence problem-solving from the research data were categorized based on theory into Frame Problem and Information Acquisition Problem. The Frame Problem involves challenges in defining the problem and goal state, as well as the difficulty of determining all possible states. The Information Acquisition Problem encompasses aspects of information modeling and reliability. Solutions proposed in the literature included various learning strategies, such as transfer learning, active learning, and experience-based learning models, as well as the modeling and measurement of search space. Across all research articles the proposed solutions varied, but a common theme was the utilization of humans as a trainer, validator of data, or problem structurer.

As a conclusion of the study is a framework where the identified problems from the material are related to the cognitive process of human problem-solving. This study does not provide an exhaustive review of artificial intelligence problem-solving and potential solutions, but it offers insights into future developments and organizes the research field from a cognitive science perspective. Based on the study, it appears that an extensive solution for enhancing artificial intelligence problem-solving capabilities has not yet been developed. Further research could investigate if the findings can be generalized with a more comprehensive dataset or explore the possibilities and weaknesses of the identified solutions.

Keywords: Intelligence, artificial intelligence, complex problem solving, cognitive science, systematic literature review

KUVIOT

Kuvio 1. Älykkyyden määritelmiä	11
Kuvio 2. Tekoälyn määritelmiä	14
Kuvio 3. Tekoälyn taksonomia	19
Kuvio 4. Ongelman visuaalinen kuvaus Smith (1991) mukaan.....	24
Kuvio 5. Ongelman kognitiivinen prosessi Wang ja Chiew (2010) mukaan	30
Kuvio 7. PRISMA-mallia mukaileva kuvaus hakutulosten seulonnasta	42
Kuvio 8. PRISMA-mallia mukaileva kuvaus toisen kierroksen hakutulosten seulonnasta	43
Kuvio 9. Tutkimusaineiston ja teorian pohjalta laadittu tulkinta tekoälyn haasteista ja mahdollisista ratkaisuista kohti ihmisen kaltaista ongelmanratkaisua.	65

LISTAUKSET

Listaus 1 Ongelmaratkaisuprosessin funktio	31
--	----

TAULUKOT

Taulukko 1. Hakutermit	39
Taulukko 2. Hakutulosten perustiedot.....	41
Taulukko 3. Tiivistelmä kirjallisuuskatsauksen keskeisistä artikkeleista	53
Taulukko 4. Aineistosta tunnistetut tekoälyn ongelmat.....	55
Taulukko 5. Aineistosta tunnistetut ratkaisuehdotukset.....	59

SISÄLLYS

1	JOHDANTO.....	7
1.1	Tutkimusongelma.....	8
1.2	Tutkimusehdotus.....	8
1.3	Tutkimusasetelma.....	9
2	ÄLYKKYYS JA ONGELMANRATKAISU	10
2.1	Ihminen – universumin älykkäin olento	10
2.2	Tekoäly	12
2.2.1	Tekoälyn kehitys	14
2.2.2	Tekoälyn menetelmät	18
2.2.3	Yleinen tekoäly	21
2.3	Ongelmanratkaisu monimutkaisessa ja muuttuvassa ympäristössä.....	23
2.4	Ihminen ongelmanratkaisijana	25
2.4.1	Gestalt-psykologit ja älynväläys	26
2.4.2	Hakuun perustuva ongelmanratkaisu	26
2.4.3	Analoginen ongelmanratkaisu	27
2.4.4	Yksilölliset erot ongelmanratkaisukyvyssä	27
2.4.5	Eri teorioita yhdistävät mallit.....	28
2.5	Tekoäly ongelmanratkaisijana	30
2.5.1	Ongelmat tekoälyn ongelmanratkaisukyvyssä	32
3	TUTKIMUKSEN TOTEUTUS.....	37
3.1	Systemaattinen kirjallisuuskatsaus	37
3.2	Tietokannan valinta	38
3.3	Hakutermien määrittely	38
3.4	Seulontakriteerit.....	40
3.5	Hakutulosten seulonta	41
3.6	Aineiston analyysi	43
3.7	Validiteetti.....	44
3.7.1	Suunnittelun validiteetti.....	44
3.7.2	Toteutuksen validiteetti.....	45
3.7.3	Tiedon käsittelyn ja analysoinnin validiteetti	47
3.7.4	Tutkimuksen tulosten käsittelyn validiteetti	49
4	TULOKSET JA SYNTEESI	51
4.1	Kirjallisuuskatsauksen tiivistelmä ja yleiset huomiot	51
4.2	Keskeiset ongelmat tekoälyn ongelmanratkaisukyvyssä	55
4.2.1	Frame-ongelma	55
4.2.2	Tiedon hankinnan ongelma	57
4.3	Ratkaisuja tekoälyn ongelmanratkaisukyvyyn kehittämiseksi.....	58
4.3.1	Ongelman selkiyttäminen	59
4.3.2	Mahdollisten tilojen määrän hallinta	60
4.3.3	Tiedon mallintaminen	61
4.3.4	Tiedon luotettavuuden vahvistaminen.....	63

5	JOHTOPÄÄTÖKSET	64
5.1	Keskeisiä ongelmia tekoälyn ongelmanratkaisukyvyssä	66
5.2	Ratkaisuja tekoälyn ongelmanratkaisukyvyn kehittämiseksi.....	67
5.3	Tutkimustulosten merkitys	68
5.4	Mahdollisia jatkotutkimuksia	69
	LÄHTEET.....	71
	LIITE 1 ESIMERKKEJÄ HAKUTERMIEN TESTAUKSESTA	76
	LIITE 2 HAKUTULOSTEN TIETEELLISEN LAADUN ARVIOINTI.....	80
	LIITE 3 ENSIMMÄINEN HAKUTULOS	81
	LIITE 4 TOINEN HAKUTULOS	84

KÄSITEHAKEMISTO

Järjestelmä = Tietojen ja käsittelysääntöjen kokonaisuus, joka pystyy suorittamaan tietojenkäsittelykokonaisuuksia toistuen samanlaisina. Suorittaa tavoitteiden mukaiset toiminnot annettujen syötteiden pohjalta.

Agentti = Tekoälyjärjestelmä tai ohjelmisto, joka kykenee itsenäiseen toimintaan ja päätöksentekoon tietyn tehtävän tai tavoitteen saavuttamiseksi.

Algoritmi = Ohjelmointiratkaisu tai sarja käskyjä, joiden perusteella jokin toiminto tapahtuu tietyllä tavalla.

Funktio = Täsmällinen sääntö tai kaava, jonka pohjalta ohjelma muuttaa syötteen tulosteeksi.

Syöte = Komento tai tieto, jonka pohjalta tietokoneohjelma tekee määriteltyjä asioita.

Tuloste = Tietokoneohjelman tuottama lopputulos annettuun tehtävään.

1 JOHDANTO

Vuoden 2022 lopussa OpenAI julkaisi yleisiin kielimalleihin pohjautuvan Chat GPT:n ja sai aikaan suuren mediakohun, kuten näistä esimerkeistä voi havaita:

Nyt puhuu tekoäly: Nämä työpaikat tapan ensin - Ihmisiantuntijat samoilla linjoilla - Talouselämä 9.4.2023

Pormestari saattaa haastaa Chat GPT:n kehittäjän oikeuteen tekoälyn väitteistä - Helsingin Sanomat 5.4.2023

Digitaalisia kaksoisolentoja ja jääkarhuja Kauppatorilla - näin ronskisti uusin tekoäly kykenee petkuttamaan sinuakin - MTV3 23.3.2023

Kohu nousi jälleen keväällä 2023 GPT-4:n julkaisun myötä, kun sen hehkutettiin olevan hätkähdyttävän lähellä ihmisen kaltaista suorituskäkyä ja varhainen versio yleisestä tekoälystä (Bubeck ym., 2023). Helsingin yliopiston tietojenkäsittelytieteen professori Hannu Toivonen kuitenkin kumoaa yleistä tekoälyä koskevat väitteet Helsingin Sanomien 19.2.2023 artikkelissa. Toivosen mukaan julistukset yleisestä tekoälystä ovat liioittelua ja sellaista tekoälyä, joka olisi kaikissa tehtävissä ihmisen veroinen, ei ole vielä kehitetty ja on epäselvää onnistuutaanko sitä koskaan kehittämäänkään (Storås, 2023). Myös uusimmissa tutkimuksissa on näyttöä sen osalta, että edes ChatGPT ei kykene tunnistamaan minikälaista maalaisjärkeä tiettyyn kysymykseen vastaaminen edellyttää (Bian ym., 2023) tai että looginen päättely on edelleen tietyiltä osin haastavaa mm. ChatGPT:lle ja GPT-4:lle (Liu ym., 2023).

Tekoälytutkimuksessa on alan alkuvaiheista lähtien pohdittu ja tehty arvioita siitä, missä vaiheessa tekoäly tulee saavuttamaan ihmisen tasoisen älykkyyden tai yleisen älykkyyden, jota voidaan verrata ihmisen älykkyyteen. Vuonna 2016 toteutetussa tutkimuksessa 50 % tekoälytutkijoista oli sitä mieltä, että tekoäly tulee suoriutumaan ihmistä paremmin 45 vuoden sisällä ja 10 % piti mahdollisena, että tämä tapahtuu yhdeksässä vuodessa (Grace ym., 2018). Russell (2020) toteaa, että *tekoälyn kehityksen kannalta ongelmana ei ole laitteisto, vaan ohjelmistolliset ratkaisut; emme vielä tiedä miten koneista saataisiin todella älykkäitä, vaikka ne olisivat universumin kokoisia.*

Viime vuosien tutkimuksissa on pyritty tutkimaan ja analysoimaan, mikä on puuttuva tekijä ihmisen ja tekoälyn välisessä älykkyydessä. Esimerkiksi Hoffmann (2022) sanoo, että uuden tyyppiset tekoälyt, kuten sosiaaliset robotit ja kognitiiviset avustajat herättävät tarpeen saada uusia vastauksia sen osalta, mitä voidaan pitää mielekkäänä vertailupohjana toisenlaiselle älykkyydelle. Ja vastaavasti Thagardin (2005) mukaan koneista voidaan tehdä älykkäämpiä pohdimalla, mikä tekee ihmisistä älykkäitä.

1.1 Tutkimusongelma

Yhtenä ihmisen korkeamman ajattelun muotona voidaan Thagardin mukaan pitää ongelmanratkaisutaitoa. Myös Jonassen (2000) sekä Wang ja Chiew (2010) mielestä yksi ihmisen olennaisimmista kognitiivisista prosesseista on ongelmanratkaisutaito. Leggin ja Hutterin (2007) näkemyksen mukaan tekoälyllä tulisi olla valmiudet kehittyä minkä tahansa ongelman ratkaisijaksi, jotta se kykenisi ihmisen kaltaiseen älykkääseen toimintaan. Tekoälyn ongelmanratkaisupotentiaalin kartoittamiseksi tässä tutkimuksessa on tarkoitus selvittää tekoälyn kykyä ja mahdollisia kehitysnäkymiä ongelmanratkaisuun monimutkaisissa ja jäsentymättömissä ongelmanratkaisutilanteissa. Tutkimuskysymykset ovat:

- 1) Minkälaisia keskeisiä ongelmia tekoälyn ongelmanratkaisukyvyyn osalta on kirjallisuudessa tunnistettu suhteessa ihmisen ongelmanratkaisukykyyn?
- 2) Minkälaisia ratkaisuehdotuksia viime vuosina on esitetty sille, että tekoäly kykenisi ongelmanratkaisuun monimutkaisissa ja jäsentymättömissä tilanteissa?

1.2 Tutkimusehdotus

Vaikka tekoälyn kehityksessä on tapahtunut merkittäviä kehitysaskelia viime vuosina, monimutkaisessa ja jäsentymättömässä ympäristössä ihminen vaikuttaa edelleen olevan edellä ongelmanratkaisutaidossa. Valtaosa tämän päivän tekoälysystemeistä on suunniteltu toimimaan rajatuissa ja ennalta määriteltyissä tilanteissa (Jiang ym., 2022). Toisaalta viime aikoina julkaistut tekoälymallit toimivat myös laajemmissa tehtäväkokonaisuuksissa (Maslej ym., 2023) ja ne pystyvät pääsemään hyvin lähelle ihmisen kaltaista suoritusta (Bubeck ym., 2023). Laajemmissa ja monimutkaisemmissa ympäristöissä toimivan tekoälyn osalta haasteena on silti tavoitteiden ja seurausten välisten päättelyketjujen monimutkaisuus. Lisäksi ihmisiin ja ympäristöön liittyvä epäohdonmukainen toiminta aiheuttaa haasteita tekoälyn ongelmanratkaisulle (Russell, 2020).

Tekoälyn ongelmanratkaisukyvyyn ymmärtäminen on tärkeää, jotta järjestelmistä voidaan kehittää luotettavia, turvallisia, toimivia ja tuloksellisia. Floridin (2020) mielestä on ilmeistä, että tarvitsemme kaiken suunniteltavissa, kehitettävissä ja käytettävissä olevan laadukkaan teknologian ratkaistaksemme maailmanlaajuisia haasteita kuten ilmastonlämpeneminen ja pakolaisuus.

Nykyisen tekoälyn hyödyntämisen haasteena ovat muun muassa epäselvät päätösprosessit sekä tiedon hankinnan ja käsittelyn virheet. Tutkimalla vaihtoehtoja tekoälyn ongelmanratkaisukyvyyn kehittämiseksi voidaan kehittää paremmin toimivia järjestelmiä, joita voidaan turvallisesti hyödyntää eri tieteenaloilla. Tai kuten Marcus ja Davis (2019) sanovat: tekoälyn tulee kehittyä merkittävästi, jotta voimme luottaa siihen meidän ongelmiemme ratkaisemisessa.

1.3 Tutkimusasetelma

Tutkimuksen aineistonkeruumenetelmänä on systemaattinen kirjallisuuskatsaus. Kirjallisuuskatsauksen etuna on, että sen avulla saadaan tehokkaasti koottua yhteen viimeaikaista tutkimustietoa, etenkin kun tekoälyn osalta tutkimuksia julkaistaan tällä hetkellä valtavasti. Tekoälytutkimuksen osalta julkaisujen määrä on yli tuplaantunut vuoden 2010 jälkeen; 2010 noin 200,000 julkaisua ja melkein 500,000 vuonna 2021 (Maslej ym., 2023). Snyder (2019) kuvaa vastavasti liiketoimintaa koskevaa tutkimuskenttää siten, että ajankohtaisen tutkimustiedon osalta on vaikea pysyä ajan tasalla tai arvioida minkälaista kollektiivista näyttöä joltakin tietyltä liiketoiminnan alueelta on saatavilla. (Snyder, 2019.) Mielestäni samat perustelut pätevät myös tekoälyn tutkimuskenttään. Uusien julkaisujen määrän, tutkimuskentän poikkitieteellisyyden ja kollektiivisen näytön arvioimisen näkökulmasta kirjallisuuskatsaus on toimiva tutkimusmenetelmä koota yhteen ja tiivistää alan viimeisintä tutkimustietoa. Kirjallisuuskatsaus on perusteltu menetelmä myös siitä näkökulmasta, että alan keskeisistä käsitteistä ei ole vakiintuneita määritelmiä.

Tarkastelen tässä tutkimuksessa tutkimusongelmaa ensisijaisesti kognitiotieteellisestä näkökulmasta eli ihminen toimii vertailukohtana tekoälyn ongelmanratkaisukyvyille. Näkökulmani on kognitiotieteellinen, koska tutkimukseni on kyseisen alan opintojen päättötyö. Toisaalta ihminen on luonteva vertailupohja myös historiallisesta näkökulmasta: tekoäly ja kognitiotiede ovat kehittyneet rinnakkain, kun on pyritty ymmärtämään ja mallintamaan ihmisen älykkyyttä. Tutkimuskysymystä tarkastellaan ns. ylätasolla, eli tarkoituksen on löytää yleiseen älykkyyteen (kappale 2.2.3) verrattavaa tietoa siitä, miten tekoälyn ongelmanratkaisua voidaan kehittää. Tekoälyn teknisiä toteutusratkaisuja en tässä tutkimuksessa tarkastele menetelmälukua laajemmin.

2 ÄLYKKYYS JA ONGELMANRATKAISU

2.1 Ihminen - universumin älykkäin olento

Ihmiset pitävät itseään älykkäimpinä koko universumissa koskaan havaituista olennoista (Korteling ym., 2021). Toistaiseksi tämä lienee myös paikkansa pitävää, sillä ihmistä korkeammalla älykkyydellä varustettua oliota ei vielä ole tunnistettu. Vuonna 1994 52 tutkijan joukko julkaisi Wall Street Journalissa lausunnon, jossa he pyrkivät määrittelemään älykkyyden perusolemuksen viimeaikaisten tutkimusten valossa. Kyseisen julkilausuman mukaan älykkyys on yleinen mentaalinen kyky, joka sisältää muun muassa päättelyn, suunnittelun, ongelmanratkaisun, abstraktin ajattelun, monimutkaisten ideoiden ymmärryksen sekä nopean oppimisen ja kokemuksesta oppimisen taidon. (Gottfredson, 1997.)

Conwayn ja Kovacsin (2015) mukaan älykkyys on kognitiotieteen historiassa yksi eniten tutkituista käsitteistä. Kognitiotiede tutkii ja pyrkii ymmärtämään ihmisen mieltä, ajattelua ja älykkyyttä. Tavoitteena on muun muassa älykkyyden periaatteiden ymmärtäminen, jotta saisimme paremman käsityksen mielen toiminnasta ja oppimisesta (Uddin, 2019). Tutkimuksellisista ponnisteluista huolimatta, vielääkään ei ole kyetty muodostamaan yksiselitteistä määritelmää, mallia tai teoriaa ihmisen älykkyydelle. (Conway & Kovacs, 2015.)

Conway ja Kovacs (2015) kuvaavat älykkyyttä koskevissa varhaisissa teorioissa olleen olennaisia eroja, mutta ne kaikki tunnistivat päättelyn, oppimisen, ongelmanratkaisun ja ymmärryksen merkityksen osana ihmisen älykkyyttä. Nämä ovat keskeisiä näkökulmia myös nykyisessä älykkyyden tarkastelussa. Thagardin (2005) mukaan jokainen vähänkin älykäs ihminen kykenee suunnittelemaan, tekemään päätöksiä ja luomaan selityksiä.

Psykometriset mallit hallitsivat pitkään älykkyyden tutkimusta Spearmanin yleisälykkyyden teorian (*general intelligence theory*) pohjalta. Myöhemmin tätä ajatusmallia haastettiin kehittämällä erilaisia moniälykkyysteorioita, joissa ihmisillä on teoriasta riippuen eri määrä älykkyyden lajeja. Nämäkin teoriat saivat kuitenkin osakseen kritiikkiä ja nykyisin hyväksytäänkin ajatus psykometrisen

yleisälykkyyden (g-faktori) olemassaolosta, mutta samalla tunnustetaan, että on olemassa tekijöitä, jotka vaikuttavat yleisälykkyyteen. (Conway & Kovacs, 2015.) Bubeck ym. (2023) sanovat, että älykkyyden ei rajoitu johonkin tiettyyn alueeseen tai tehtävään, vaan on laaja joukko kognitiivisia kykyjä ja taitoja.

Russell on sitä mieltä, että älykkyyttä ei määritellä AO testien tai Turingin testin pohjalta, vaan se muodostuu havainnoin, tavoitteiden ja toiminnan välisestä suhteesta (*what we perceive, what we want, and what we do*). Oliota voidaan pitää älykkäänä, jos se kykenee saavuttamaan haluamansa huomioiden havaitsemansa. (Russell, 2020.) Legg ja Hutter (2007) ovat Russellin kanssa samoilla linjoilla omassa määritelmässään, jonka he ovat muotoilleet tarkasteltuaan älykkyyden määritelmiä tietosanakirjojen, psykologien ja tekoälytutkijoiden pohjalta. Tutkittuaan erilaisia älykkyyden määritelmiä Legg ja Hutter ovat tunnustaneet niistä kolme yhdistävää piirrettä: ominaisuus toimittaessa suhteessa ympäristöön, tavoitteen tai määränpään onnistunut saavuttaminen sekä kyky mukautua erilaisiin tavoitteisiin ja ympäristöihin. Älykkyyden määritelmiä yhdistävien piirteiden pohjalta Legg ja Hutter määrittelevät älykkyyden mittaavan toimijan kykyä saavuttaa tavoitteita monenlaisissa ympäristöissä. (Legg & Hutter, 2007.)

Olen kuvioon 1 koonnut muutamia tässä kappaleessa esiteltyjä älykkyyden määritelmiä. Itse ajattelisin Russellin (2020) sekä Leggin ja Hutterin (2007) määritelmän olevan kattavampi, kuin pelkkään ajattelulliseen tai mentaaliseen näkökulmaan pohjautuvat määritelmät. Pidän tärkeänä, että älykkyyden määritelmään sisältyy kyky suunnitella ja tehdä päätöksiä, mutta myös kyky mukautua erilaisiin ongelmantilanteisiin sekä suunnitelmallinen toiminta kohti määränpäättä. Omasta mielestäni kattava älykkyyden määritelmä sisältäisi näkökulmia jokaisen kuvion 1 määritelmistä ja voisi olla jotakin tämän kaltaista: Älykkyyden on joukko mentaalisia kykyjä, joita vaaditaan toiminnan suunnittelemiseksi ja päätösten tekemiseksi määriteltyjen tavoitteiden saavuttamiseksi monenlaisissa ympäristöissä.



Kuvio 1. Älykkyyden määritelmiä

2.2 Tekoäly

Tekoälyn medianäkyvyyden myötä useimmilla ihmisillä lienee jonkinlainen käsitys siitä, mitä tekoälyllä tarkoitetaan. Harva tuskin pystyy kuitenkaan tarkemmin määrittelemään, mitä tekoäly oikeastaan on. Toisaalta tekoälyn määrittely ei ole ihan yksinkertaista tutkijoillekaan, kuten Hoffmann (2022) sanoo: mikä tahansa määritelmä, joka huomioisi kaikki tekoälyn eri ulottuvuudet olisi todennäköisesti niin kaikenkattava, että se olisi lopulta merkityksetön.

Ensimmäisen kerran tekoäly mainittiin tieteellisessä kontekstissa vuonna 1955, jolloin McCarthy, Minsky, Rochester ja Shannon (2006)¹ tekivät ehdotuksen Dartmouthin seminaarista. He kutsuivat tutkijoita koolle konferenssiin, jonka tarkoituksena olisi kehittää kone, jonka toimintaa kutsuttaisiin älykkääksi, jos ihminen käyttäytyisi vastaavalla tavalla. (McCarthy ym., 2006.) Myöhemmin McCarthy määritteli Hayesin (1979) kanssa, että voidakseen toimia älykkäästi, koneen tulee kyetä tulkitsemaan syötteitä suhteessa yleisiin representaatioihin maailmasta. Myös Dennett (1984) määrittelee älykkyyden muodostuvan toiminnan ja ympäristön vuorovaikutuksesta. Dennettin mukaan älykkyys on tiedon mielekästä hyödyntämistä. Tekoälyn tulisi kyetä suunnittelemaan toimintaansa ja hyödyntämään hyvin valittuja toimintamalleja suhteessa maailmaan, jossa se toimii. (Dennett, 1984.)

Bubeck ym. (2023) mukaan tekoälytutkijoiden pitkäaikainen ja kunnianhimoinen tavoite on kehittää keinotekoinen järjestelmä, joka kykenee osoittamaan 52 tutkijan (ks. kappale 2.1) määritelmän mukaista yleisälykkyyttä. Myös Norvig ja Russellin (2010) mielestä tekoälytutkimuksen pyrkimyksenä ei ole pelkästään älykkäiden kokonaisuuksien ymmärtäminen vaan myös sellaisen luominen. Goertzelin ja Wangin (2007) mielestä täydellisesti toimiva yleinen tekoäly² tulee hyvin todennäköisesti muistuttamaan monella tapaa ihmisen aivoja tai ajattelua. He eivät kuitenkaan halua määritellä yleistä tekoälyä tämän pohjalta, vaan yleinen tekoäly voidaan saavuttaa hyvin erilaisten teoreettisten ja teknisten lähestymistapojen kautta. Goertzel ja Wang lisäävät, että vielä on liian varhaista todeta, mikä älykkyyden tulkinta on lopulta ”oikea”. Tulkinta tulee heidän mielestään olla jokaisella tutkijalla itsellään. (Goertzel & Wang, 2007.)

Hoffmann (2022) tuo esille, että älykkyyden lisäksi tekoälyn kehitystä voidaan tarkastella myös toiminnan näkökulmasta: onko tavoitteena saavuttaa vastaavuus ihmisen toimintakyvylle vai rationaalisuudelle. Tarkoitetaanko käytännössä ajatteluun kykeneviä järjestelmiä vai toimintaan perustuvia järjestelmiä,

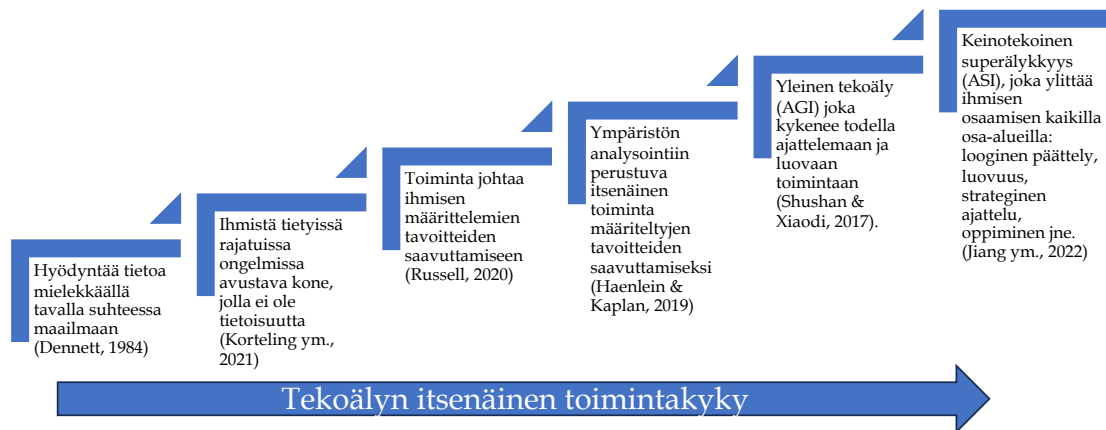
¹ Alkuperäistä konferenssiehdotusta säilytetään Dartmouth College and Stanford Universityn arkistossa, joten olen hyödyntänyt lähteenäni 2006 AI Magazinessa julkaistuja ehdotuksen pääkohtia.

² Yleisen tekoälyn tarkempaan määrittelyyn palaan tarkemmin kappaleessa 2.2.3.

kuten robotiikkaa. (Hoffmann, 2022; Norvig & Russell, 2010) Haenlein ja Kaplan (2019) ovatkin sisällyttäneet tekoälyn määritelmään myös toiminnan näkökulman. Heidän mukaansa tekoälyllä viitataan järjestelmiin, jotka kykenevät ympäristöönsä analysoimalla tekemään itsenäisiä toimia tiettyjen tavoitteiden saavuttamiseksi. Haenlein ja Kaplan lisäävät myös, että tekoälyjärjestelmät voivat olla täysin ohjelmistopohjaisia, mutta myös laitteisiin sisäänrakennettuja. (Haenlein & Kaplan, 2019.) Myös Russell (2020) määrittelee koneiden älykkyyden olevan suhteessa siihen, kuinka hyvin toiminta johtaa asetettujen tavoitteiden saavuttamiseen. Russell kuitenkin jatkaa, että koneilla ei ole omia tavoitteita, eikä meillä ole luotettavaa tapaa arvioida, ovatko koneiden tavoitteet samat kuin meidän tavoitteemme. Russellin mielestä tekoälyn luonnehdinnan tulisi siis itseasiassa olla: koneet ovat hyödyllisiä siinä laajuudessa, kuin niiden toiminnan voidaan ajatella saavuttavan meidän määrittelemämme tavoitteet. (Russell, 2020.)

Edellä kuvattujen tekoälyn määritelmien pohjalta olen itse taipuvainen ajattelemaan Hoffmanin (2022) tapaan, että McCarthy ym. (2016) alkuperäinen määritelmä siitä, että tekoälyä voidaan pitää älykkäänä, jos vastaavaa toimintaa ihmisen kohdalla pidettäisiin älykkäänä, on varsin onnistunut. Toisaalta myös Russellin (2020) näkemys siitä, että älykkyys muodostuu sen perusteella, saavutetaanko toiminnalle asetetut tavoitteet huomioiden ympäristöstä tehdyt havainnot, on mielestäni hyvä. Toisaalta Russellin määritelmä koneiden hyödyllisyydestä suhteessa ihmisen määrittelemiin tavoitteisiin sulkee mielestäni jossain määrin pois ajatuksen yleisestä tekoälystä.

Lopuksi todettakoon Hoffmannia (2022) lainaten, että niin moninaisia tutkimuskohteita sisältävän tutkimusalan, kuin tekoäly, kuvaaminen kattavasti on lähes ylitsempääsemätön tehtävä. Itse olen kuvioon 2 hahmotellut kirjallisten lähteiden ja oman tulkintani pohjalta tiivistelmän tekoälyn määritelmästä. Kuvasta puuttuu muun muassa tietoisuutta ja neurotieteellistä lähestymistä koskevat näkökulmat, jotka olen rajannut pois myös gradustani. Kuvan pääasiallinen tarkoitus on kuvastaa, miten laajasti tekoäly voidaan ymmärtää ja millä tasolla nykyisen tekoälyn voidaan ajatella olevan. Tekoälyn määritelmien lisäksi olen sisällyttänyt kuvaan oman tulkintani siitä, että tekoälyn älykkyyden lisääntyessä myös sen itsenäinen toimintakyky kasvaa. Oman tulkintani mukaan nykyinen tekoäly on kuvan mukaisista tasoista jossakin ensimmäisen ja kolmannen portaan väli-
maastossa riippuen tarkastelukulmasta.



Kuvio 2. Tekoälyn määritelmiä

2.2.1 Tekoälyn kehitys

Vuosisatojen ajan ihminen on kuvitellut, haaveillut ja sepittänyt taruja älykkäistä koneista. Keinotekoinen älykkyys on ollut niin filosofien, tutkijoiden kuin kirjailijoidenkin mielenkiinnon kohteena. Ensimmäiset todelliset tekoälyn kehitysaskeleet on kuitenkin otettu vasta 1900-luvulla. (Buchanan, 2005.) Norvigin ja Russellin (2010) mukaan työskentely tekoälyn parissa alkoi toisen maailmansodan jälkeen ja *artificial intelligence* termi on keksitty 1956. Haenleinin ja Kaplanin (2019) mielestä tekoälyn historiassa merkittävänä inspiroijana voidaan pitää amerikkalaisen tieteiskirjailija Isaac Asimovin novellia Runaround, jossa Asimov esitteli robotiikan kolme pääsääntöä. Runaround julkaistiin vuonna 1942 ja sen jälkeen tekoälytutkimuksessa ovat vuorotelleet niin onnistumiset kuin takaiskutkin; tekoälyn on vuoroin uskottu olevan aivan muutaman vuoden kehityksen päässä ja vuoroin epäilty saavuttamattomaksi. (Haenlein & Kaplan, 2019.)

Seuraavissa kappaleissa kuvaan lyhyesti muutamia tekoälyn historian tärkeimpiä kehitysvaiheita. Historiakatsauksen osalta on hyvä muistaa, että tekemäni nostot perustuvat tiettyihin julkisuutta saaneisiin saavutuksiin, mutta niidenkin takana on myös muita eri tieteenaloilla tehtyjä tutkimuksia. Kuten Buchanan (2005) sanoo, varhaistenkin merkkipaalujen takana on tutkimusta muun muassa ongelmanratkaisun, oppimisen, tiedon representaatioiden ja assosiattiivisen muistin osalta. Myös Russell (2020) korostaa, että mediasta välittyvä kuva tekoälyn läpimurroista ei tuo näkyväksi vuosikymmenien aikana tehtyä tutkimus- ja kehitystyötä.

Haenlein ja Kaplan (2019) ovat jakaneet omassa tekoälyn historiaa, nykytilaa ja tulevaisuutta kuvaavassa artikkelissaan tekoälyn kehityksen neljään vaiheeseen: kevät, kesä ja talvi, syksy ja nykyhetki. Floridi (2020) huomauttaa, että vuodenaika metaforan käytön ongelmana on, että vuodenaikat seuraavat syklistisesti toisiaan ja talvi tulee aina uudelleen. Itse olen jakanut kehityksen viiteen vaiheeseen, joissa onnistumisten kausia seuraavat tekoälytalvet ja viimeisenä vaiheena on 2010-luvulta nykypäivään sijoittuva ajanjakso, jolloin on tapahtunut useita merkittäviä edistysaskelia.

Alkuinnostus ja onnistumisia 1956–1974

Turingin 1940 ja 1950-luvuilla tietokoneiden parissa tekemää työtä pidetään tekoälytutkimuksessa yhtenä keskeisistä historiallisista murroksista. (Buchanan, 2006). Myös Norvig ja Russell (2010) mainitsevat, että useampi tutkija työskenteli tekoälytutkimukseen liittyvien teemojen parissa, mutta Alan Turingin visiota voidaan pitää kaikkien vaikuttavimpana. Vuonna 1950 julkaistussa artikkelissa *Computing Machinery and Intelligence* Turing esitteli Turingin testin, koneoppimisen, geneettisen algoritmin ja vahvistusoppimisen. (Norvig & Russell, 2010.)

Käsite tekoäly kehitettiin vuonna 1956 Dartmouthissa pidetyssä kahdeksanviikkoisessa konferenssissa *Dartmouth Summer Research Project on Artificial Intelligence* (DSRPAI) (Haenlein & Kaplan, 2019). Konferenssin tavoitteena oli selvittää, miten koneet saataisiin tuottamaan kieltä, muodostamaan käsitteitä ja ajatuksia, ratkomaan ongelmia ihmisten tavoin sekä kehittämään itse itseään. Ajatuksena oli, että mikä tahansa oppimisen tai älykkyyden muoto voidaan kuvata niin tarkasti, että sitä simuloimaan on mahdollista kehittää kone. (McCarthy ym., 2006.)

Vaikka Dartmouthin konferenssi ei tuottanut toivottua läpimurtoa, siihen osallistuneet tutkijat ja heidän kollegansa hallitsivat alaa seuraavat 20 vuotta: Esimerkiksi Newell ja Simon kehittivät *General Problem Solver:n* (GPS) 1957, McCarthy määritteli LISP-ohjelmointikielen 1958 ja Samuel kirjoitti 1952–1956 itseoppivan peliohjelman tammipelille. (Norvig & Russell, 2010.)

Edellä mainittujen sekä muiden (mm. ELIZA) onnistuneiden projektien ansiosta tekoälytutkimus sai hyvin rahoitusta, mikä johti myös tutkimusprojektien lisääntymiseen. Vuonna 1970 Marvin Minsky oli niin optimistinen, että sanoi tavallista ihmistä vastaavan yleisälykkyyden olevan kehitettävissä kahdeksassa tai kymmenessä vuodessa. (Haenlein & Kaplan, 2019.) Myös Herbert Simon sortui 1957 yltioptimismiin sanoessaan koneiden kykenevän lähitulevaisuudessa yhtä laajaan ongelmien käsittelyyn, kuin ihminen (Norvig & Russell, 2010). Epärealististen odotusten ja saavutusten välisen epäsuhdan myötä usko tekoälyn kehitykseen alkoi rakoilla ja rahoituskanavat tyrehtyivät, mikä johti ensimmäiseen tekoälytalveen (Haenlein & Kaplan, 2019; Norvig & Russell, 2010).

Ensimmäinen tekoälytalvi 1974–1980

Floridi (2020) määrittelee tekoälytalven (*AI winter*) tarkoittavan vaihetta, jolloin tekoälyn ympärillä pyörivä teknologian, liike-elämän ja median into laantuu ja havahdutaan teknologian todellisiin mahdollisuuksiin. Ensimmäinen tekoälytalvi seurasi, kun tekoälyn kehitys ei vastannutkaan suuriin odotuksiin ja kehityksessä tuli vastaan taloudellisia takaiskuja. Suurta innostusta seurasi pettymys. (Jiang ym., 2022.) Vuonna 1973 brittiläinen matemaatikko James Lighthill julkaisi raportin, jossa hän kyseenalaisti tekoälytutkijoiden optimistisia näkemyksiä (Haenlein & Kaplan, 2019). Yksi Lighthillin raportin keskeisimmistä kritiikin aiheista oli epäonnistuminen koneiden laskennallisen tehon kasvattamisessa (Norvig & Russell, 2010). Lighthill totesi myös, että koneet tulevat saavuttamaan ainoastaan kokeneen amatöörin tason sellaisissa peleissä, kuin shakki ja maalaisjärkeen perustuva päättely tulee aina olemaan saavuttamattomissa. Tämän

seurauksena lopetettiin tekoälytutkimuksen rahoitus valtaosassa Brittiyliopistoja ja sitten myös Yhdysvalloissa. (Haenlein & Kaplan, 2019.)

Kiinnostus kaupallisiin mahdollisuuksiin kasvaa 1980–1987

Osassa lähdeaineistoani tekoälytalveksi katsotaan koko ajanjakso vuodesta 1974 aina 1990-luvulle asti, koska 80-luvulla ei katsota tapahtuneen mitään merkittävää kehitystä (mm. Haenlein & Kaplan, 2019). Toisaalta tekoälyteollisuuden käynnety rahoitus kasvoi vuoden 1980 muutamasta dollarista biljooniin dollareihin 1988 ja useat yritykset rakensivat muun muassa hahmontunnistusjärjestelmiä, robotteja sekä ohjelmistoja ja laitteita näihin tarkoituksiin (Norvig & Russell, 2010). Myös Jiang ym. (2022) mukaan uusi kiinnostuksen aalto nousi, kun tekoälyteknologiat lähtivät kehittämään kaupallisia sovelluksia, jotka oli toteutettu sääntöpohjaisilla lausekkeilla (*if-then*). Ihmisten puutteelliset tiedot tekoälyn rajoituksista monimutkaisten tehtävien ja järjestelmien, kuten lääketieteellisen diagnostiikan ja hahmontunnistuksen, osalta johtivat kuitenkin jälleen kerran liialliseen hehkutukseen tekoälyn ympärillä. (Jiang ym., 2022.)

Toinen tekoälytalvi 1987–1994

Jiangin ym. (2022) mukaan historiassa on ollut kaksi tekoälytalvea. Toinen tekoälytalvi seurasi, kun iso osa yrityksistä epäonnistui liioiteltujen lupaustensa lunastamisessa (Jiang ym., 2022; Norvig & Russell, 2010). Haenleinin ja Kaplanin (2019) mukaan yksi syy kehityksen puutteeseen oli käsitys siitä, että ihmisen älykkyys voitaisiin formalisoida ja rakentaa uudelleen sääntöpohjaisten lausekkeiden näkökulmasta. Myös matematiikan, sähkötekniikan ja tietotekniikan kehittymättömät työkalut, kuten algoritminen tulokulma, tietovarannot ja laskentateho hidastivat kehitystä. Sijoittajat eivät myöskään olleet tietoisia riskeistä ja aliarvoivat olennaisen tutkimuksen tekemiseen kuluvan ajan. (Jiang ym., 2022.)

Elpyminen ja uusi nousu 1994–2020

Tekoälyn kehitys lähti uuteen nousuun 1990-luvun alkupuoliskolla, mihin Jiangin ym. (2022) mukaan vaikutti useampi ratkaiseva tekijä:

1. onnistumiset koneoppimisen saralla
2. tutkijoiden panostus koneoppimisen kannalta oleellisten tietoa-ineistojen kehittämiseen ja kokoamiseen
3. laskentatehon merkittävä lisääntyminen ja
4. julkisen uskon ja kiinnostuksen palautuminen ihmisen ja tekoälyn välisissä kilvoitteluisa saavutettujen voittojen myötä.

Ensimmäisiä merkkipaaluja tekoälyn ja ihmisen välisessä kilpailussa oli, kun toukokuussa 1997 IBM:n Deeb Blue voitti shakin hallitsevan maailmanmestarin Garry Kasparovin (Shushan & Xiaodi, 2017). Toisaalta Deep Bluen toiminta perustui sääntöpohjaisiin lausekkeisiin, jolloin sen toimintakyky oli kapea ja

rajoittui ainoastaan vastaavalla logiikalla toimiviin tehtäviin (Haenlein & Kaplan, 2019).

Tekoölyn osalta 2000-luvulla saatiin todistaa myös uutta kiinnostusta koneoppimisen suuntaukseen (Hoffmann, 2022). Koneoppimiseen pohjautui myös vuonna 2011 paljon huomiota saavuttanut IBM:n tekoäly Watson, joka voitti kuuluisan amerikkalaisen television tietovisaohjelman "Jeopardy!". Watson ei ollut yhteydessä internetiin, mutta siinä oli neljän teratavun suuruinen tallennustila, joka sisälsi yli 200 miljoonaa sivua tietoa, mukaan lukien koko Wikipedian sisältö. (Jiang ym., 2022.)

Tekoölyn "toinen aalto" tai "uusi kevät" alkoi syväoppimiseen perustuvien ohjelmien, kuten AlphaGon myötä (Hoffmann, 2022). Vaikka pitkään uskottiin, että tietokoneet eivät tule koskaan päihittämään ihmistä Go-lautapelissa, Googlen kehittämä AlphaGo voitti vuonna 2015 maailmanmestari Lee Sedolin (Haenlein & Kaplan, 2019). AlphaGo Zero voittoa pidetään erityisen merkittävänä, koska se oppi ilman minkäänlaista opetusaineistoa pelaamaan Go-peliä ihmistä paremmin. AlphaGo Zero jopa ylitti ihmisille tuhansien vuosien aikana karttuneet tiedot ja kehitti muutamassa päivässä uusia pelistrategioita. (Jiang ym., 2022.) AlphaGon kehittämisessä hyödynnettiin joukkoa uusia strategioita, kuten syväoppimista, neuroverkkoa ja arvon arviointijärjestelmää (Shushan & Xiaodi, 2017).

Nykyhetki ja kielimallit 2020

Tekoälyä koskevat teoriat ja teknologiat ovat olleet 50-luvulta lähtien jatkuvassa kehityksessä ja parantelun kohteena (Shushan & Xiaodi, 2017). Norvig ja Russell (2010) sanovat, että tänä päivänä on paljon yleisempää lähteä kehittämään olemassa olevien teorioiden päälle, kuin ehdottaa kokonaan uusia. Vuoden 2011 paikkeilla, syväoppimisen tekniikoilla saavutettiin merkittävää edistystä tekoälykentän kolmessa keskeisessä haasteessa; puheen tunnistuksessa, visuaalisten objektien tunnistuksessa ja konekääntämisessä. Nykyään koneet jopa jossain määrin ylittävät ihmisen kyvyt näillä alueilla. (Russell, 2020.)

Kehityksen myötä kiinnostus yleistä tekoälyä kohtaan on kasvanut entisestään. Tekoäly on lehtien pääotsikoissa lähes päivittäin, tuhansia uusia yrityksiä on perustettu riskisijoituksilla, miljoonat opiskelijat ovat suorittaneet tekoölyn verkkokursseja ja alan huippuosajat tienaaavat miljoonia dollareita. Riskisijoittajien, kansallisten hallitusten ja suuryritysten tekoölyyn sijoittama pääoma on kymmeniä biljoonia dollareita vuosittain, viimeisimpien vuosien panostus on enemmän kuin kaikkina tekoälyhistorian aiempina vuosina yhteensä. (Norvig & Russell, 2010.)

Haenlein ja Kaplan (2019) kuvaavat yhtenä esimerkkinä tekoölyn arkipäiväistymisestä sitä, että termi "botti" on siirtynyt myös arkikieleen. Viimeisimpänä innostuksen aiheena ovatkin olleet vuosina 2022 ja 2023 OpenAI:n julkaisemat ChatGPT:t GPT-3 ja GPT-4. Maaliskuussa 2023 julkaistu GPT-4 on kielimalipohjainen tekoäly, joka on kehitetty ymmärtämään ja tuottamaan multimodaalista sisältöä. Sen on markkinoitu omaavan myös entistä tehokkaampia ominaisuuksia tehtävissä, jotka edellyttävät loogista päättelyä. (Liu ym., 2023.)

Tulevina vuosina tekoäly tulee Buchanan (2005) mukaan olemaan osa jokapäiväistä elämäämme internetin ja sosiaalisen median tapaan. Tämä tulee vaikuttamaan sekä henkilökohtaiseen elämäämme, mutta myös muuttamaan perustavanlaatuisesti sitä, miten yrityksissä tehdään päätöksiä ja toimitaan osakkeenomistajien kanssa. Yleisestä tekoälystä ja tulevaisuuden näkymistä sen osalta kerrotaan tarkemmin kappaleessa 2.2.3.

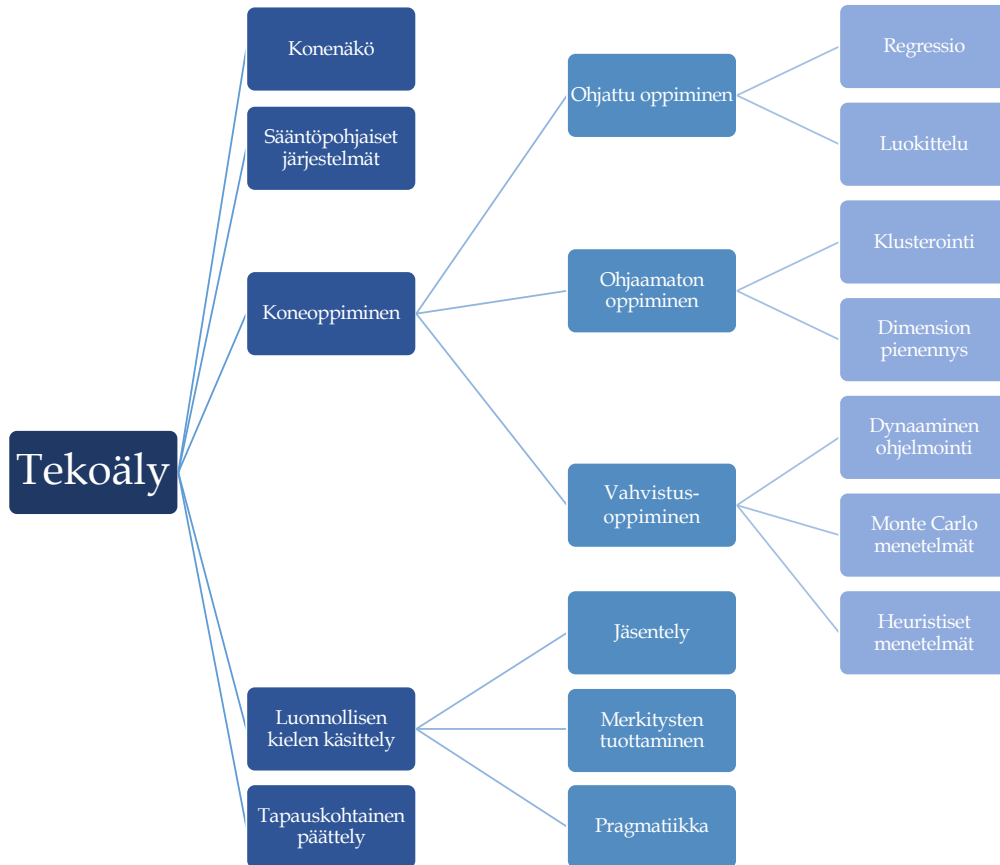
2.2.2 Tekoälyn menetelmät

Tekoälyn menetelmiä on vuosikymmenten aikana kehitetty useiden eri tieteenalojen (mm. filosofia, psykologia, matematiikka, kognitiotiede, kielitiede, neurotiede, tietojenkäsittely) teorioiden pohjalta. Jiang ym. (2022) mukaan tutkimuskentän valtavirtaa ovat konenäkö, hahmontunnistus, robotiikka ja automaatio, luonnollisen kielen käsittelymallit, tiedon etsintä ja datan louhinta sekä päätöksentekostrategiat. Tekoälyn menetelmät voidaan ryhmitellä esimerkiksi seuraavasti (Hoffmann, 2022):

- Sääntöpohjaiset järjestelmät (klassinen tekoäly)
- Koneoppiminen
- NLP eli luonnollisen kielen käsittelymallit
- CBR eli tapauskohtainen päättely
- Konenäkö

Hoffmanin (2022) ryhmittelyyn sisältyy yllä mainittujen tekniikoiden lisäksi myös tavoitepohjainen osuus, johon sisältyvät robotiikka, suunnittelu ja tietokonepelaaminen. Muita kirjallisuudesta löytyviä tekoälymenetelmien luokittelutapoja on esimerkiksi havainnointiin, toimintaan ja representaatioihin perustuva luokittelu (mm. Norvig & Russell, 2010). Koska yllä oleva Hoffmanin luokittelu on mielestäni tiivis ja selkeä, käytän sitä oman menetelmäkuvaukseni pohjana. Tiedostan, että kyseinen luokittelu ei ole kaiken kattava tai yleisesti käytetty, mutta mielestäni se on onnistunut tiivistelmä tutkimukseni kannalta keskeisistä menetelmistä. Kuvioon 3 olen laatinut Hoffmannin (2022), Hossainin ym. (2020), Mukhamedievin ym. (2022) ja Lensun (2002) pohjalta oman tulkintani tekoälyn menetelmien taksonomiasta.

Kuvailen seuraavissa kappaleissa tarkemmin koneoppimisen, luonnollisen kielen käsittelymallien sekä tapauskohtaisen päättelyn menetelmiä. Sääntöpohjaisten järjestelmien ja konenäön menetelmät rajaavat ulkopuolelle, koska ne soveltuvat muita menetelmiä huonommin monimutkaisten ongelmien ratkaisemiseen. Sääntöpohjaisten järjestelmien kohdalla ongelmana on muun muassa kaikkien mahdollisten sääntöjen ja vaiheiden määrittelyn vaikeus, kuten Jiang ym. (2022) mainitsevat. Vastaavasti konenäkö rajoittuu yleensä ennalta määriteltyjen tehtävien tai kohteiden tunnistamiseen. Koneoppimisen menetelmät kuvaavat yleisesti käytössä olevien (mm. Miao ym., 2023) alaluokkien pohjalta: ohjattu oppiminen, ohjaamaton oppiminen ja vahvistusoppimiseen.



Kuvio 3. Tekoälyn taksonomia

Ohjattu oppiminen

Ohjatun oppimisen perusajatuksena on, että agentille tuotetaan ensin opetusaineisto, jonka pohjalta se opettelee tunnistamaan syötteiden ja tulosteiden välisiä luokittelumalleja. Tekoälyllä ei siis ole ihmisen valmiiksi määrittelmiä sääntöjä, vaan se oppii muodostamaan algoritmit annetun opetusmateriaalin pohjalta. (Jiang ym., 2022.) Oppiminen tapahtuu etsimällä mahdollisten hypoteesien joukosta sellaisia funktioita, joiden avulla on mahdollista suoriutua myös koulutusaineiston ulkopuolisista uusista esimerkeistä (Norvig & Russell, 2010).

Käytännössä ohjattua oppimista hyödynnetään luokitteluun, kuten kuvien, hahmojen ja tekstin tunnistukseen. Ohjatun oppimisen mukaisia teknisiä ratkaisuja ovat muun muassa luokittelu, regressio, lineaarinen regressio, neuroverkot, tukivektorikoneet, logistinen regressio ja lineaarinen erotteluanalyysi. Ohjatun oppimisen haasteena on, että se edellyttää yleensä laajaa esimerkkiluokkiin perustuvaa opetusaineistoa. Valmiiksi luokitellun luotettavan aineiston tuottaminen on usein työlästä (Jiang ym., 2022).

Ohjaamaton oppiminen

Ohjaamattomassa oppimisessa agentti oppii tunnistamaan syötteisiin sisältyviä kaavoja (Norvig & Russell, 2010). Käytännössä tekoäly siis luokittelee ja analysoi aineistoa itse ja muodostaa sen pohjalta samankaltaisten asioiden välisiä ryhmiä. Ohjaamattoman oppimisen etuna on, että koulutusaineiston ei tarvitse olla valmiiksi lajiteltua. Ryhmittelyyn perustuvat menetelmät voivat jakaa aineiston automaattisesti ryhmiin ilman, että pääpiirteitä on tarkasti kartoitettu. (Jiang ym., 2022.) Ohjaamattomassa oppimisessa hyödynnetään esimerkiksi klusterointiin, syväoppiviin neuroverkkoihin, assosiaatioon sekä poikkeavuuksien havainnointiin perustuvia tekniikoita.

Vahvistusoppiminen

Vahvistusoppimisessa agentti oppii toimimaan vahvistavan tai rankaisevan palautteen perusteella. Tavoitteena on oppia tekemään valintoja, jotka johtavat ympäristön tai tilanteen kannalta optimaalisimpaan lopputulokseen. (Norvig & Russell, 2010). Esimerkiksi pelin voitto tai tappio ovat palautteita, joiden perusteella agentti voi arvioida toteutettujen toimintojen onnistumista. Jiang ym. (2022) mukaan hyvä esimerkki vahvistusoppimiseen perustuvasta algoritmista on AlphaGo Zero.

Vahvistusoppimiseen perustuvia lähestymistapoja ovat mm. Q-oppiminen, Monte Carlo, Brute Force, Markov-päätösprosessi ja arvioiterointi. Useilla monimutkaisilla sovellusalueilla vahvistusoppiminen on ainoa toteuttamiskelpoinen tapa saada ohjelma toimimaan laadukkaasti (Norvig & Russell, 2010).

Koneoppimisen laajennukset

Edellä mainittujen koneoppimisen menetelmien myötä tekoälyn kehityksessä on saavutettu merkittävää edistystä, mutta Miao ym. (2023) mukaan menetelmissä on havaittu ongelmia esimerkiksi puuttuvien tietojen, virheellisen luokittelun ja tilastollisten oletusten osalta. Ohjaamaton oppiminen, ohjattu oppiminen ja vahvistusoppiminen toimivat kuitenkin koneoppimisen perustana ja luovat pohjan uusien menetelmien kehittämiseksi. Viime vuosina niiden pohjalta on kehitetty joukko uusia menetelmiä, kuten aktiivinen oppiminen (active learning), siirto-oppiminen (transfer learning), elinikäinen oppiminen (lifelong learning) ja rinnakkaisoppiminen (parallel learning). (Miao ym., 2023.)

Luonnollisen kielen käsittely

Viime vuosien merkittävimpänä tekoälytutkimuksen läpimurtona pidetään suurten kielimallien (LLM) avulla saavutettua edistystä luonnollisen kielen käsittelyssä (Bubeck ym., 2023). Luonnollisen kielen käsittelymallit, kuten ChatGPT, perustuvat valtavien tietoaaineistojen pohjalta ennalta koulutettuihin kielimalleihin (Miao ym., 2023). Käytännössä kielimallit tulkitsevat ja tuottavat ihmisten käyttämää kieltä erilaisten laskennallisten mallien pohjalta (Chang ym., 2023). Laskemalla todennäköisyyksiä aikaisemman aineiston tai käytettävissä olevan

tiedon pohjalta malli arvioi, millä todennäköisyydellä sanat esiintyvät tietyssä yhteydessä ja tietyssä järjestyksessä. Esimerkiksi moniulotteisten vektoreiden avulla arvioidaan sanojen semanttista samankaltaisuutta (Edinger & Goldstone, 2022) ja N-gram malli arvioi sanojen esiintymisen todennäköisyyksiä aiemman kontekstin perusteella (Chang ym., 2023). Muita luonnollisen kielen käsittelyyn pohjautuvia malleja ovat muun muassa GPT (*Generative Pre-trained Transformer*) ja *word-embedding*.

Tekoälytutkimuksessa koneen kykyä ymmärtää ja käsitellä ihmisen tuottamaa kieltä pidetään erityisen merkittävänä siitä syystä, että kieli on oleellinen osa ihmisen kaltaista oppimista, ajattelua ja tiedon käsittelyä (mm. Norvig & Russell, 2010; Shushan & Xiaodi, 2017). Kone, joka todella ymmärtää³ ihmisen kieltä, pystyy Russellin (2020) mielestä ottamaan haltuun laajoja määriä ihmisen tuottamaa tietoa ja siten ohittamaan tiedollisesti kaikki maanpäällä liikkuneet ihmiset. Miao ym. (2023) mielestä multimodaalista ennalta koulutettua kielimallia voidaan pitää mahdollisena polkuna kohti yleistä tekoälyä.

Tapauskohtainen päättely

Tapauskohtaisen päättelyn (Case-Based Reasoning eli yleisesti CBR) ajatuksena on hyödyntää ongelmien ratkaisemisessa aiempia kokemuksia tai olemassa olevaa tietoa. Watson (1999) kuvaa CBR järjestelmän toimivan siten, että ensin aiempien tapausten kirjastosta etsitään käsillä olevan ongelman kanssa yhdenmukainen tilanne. Tämän jälkeen arvioidaan tapausten yhdenmukaisuutta ja mikäli yhdenmukaisuusvaatimukset täyttyvät, yrittää järjestelmä käyttää löytämäänsä ratkaisua ongelman ratkaisemiseen. Lopuksi ratkaisu lisätään tapauskirjastoon säilöön ja kartutetaan siten osaamista tulevia tilanteita ajatellen. (Watson, 1999.) CBR menetelmän teknisessä toteutuksessa hyödynnetään muun muassa lähimmän naapurin menetelmää, fuzzy logica, induktiota ja SQL:aa.

2.2.3 Yleinen tekoäly

Shushan ja Xiaodi (2017) määrittelevät yleisellä tekoälyn tarkoittavan sellaista konetta, joka pystyy todella ajattelemaan. Tulkitsen tämän tarkoittavan älykkyyttä, joka osoittaa myös ymmärrystä (vrt. Kiinanalaisen huoneen argumentti³). Goertzel ja Wang (2007) mukaan yleinen tekoäly eli AGI viittaa tutkimusalaan, jossa tutkitaan erikoistuneen tekoälyn sijaan yleiseen älykkyyteen perustuvia järjestelmiä. Yleisellä älykkyydellä tarkoitetaan älykkyyttä sellaisena kokonaisuutena, joka muistuttaa ihmisen ajattelua tai vastaa laajuudeltaan ihmisen ajattelukykyä. (Goertzel & Wang, 2007.) Yleisen tekoälyn tulisi siis kyetä suoriutumaan ihmistä paremmin lähes kaikissa kognitiivisissa tehtävissä (Lu ym., 2018).

Viime vuosien kehityksestä huolimatta yleinen näkemys on, että nykyinen tekoäly on valtaosin heikkoa tekoälyä eli teknologiaa, joka on keskittynyt tietyille

³ Ymmärtääkö luonnollisen kielen käsittelyyn pohjautuva kone kieltä, jos sille pystyy syöttämään tekstiä ja se kykenee tuottamaan lauseita? John Searlen Kiinanalaisen huoneen argumentin mukaan sääntöjen ja ohjeiden mukaisen sisällön tuottaminen ei vielä ole osoitus ymmärryksestä.

osa-alueille kuten kuvien tai puheen tunnistukseen tai keskusteluihin vastaamiseen. (Lu ym., 2018). Goertzel ja Wang (2007) muistuttavat, että koneoppimiseen perustuvat algoritmit eivät kuulu yleisen tekoälyn alueelle, koska oppiminen tapahtuu eristyksessä ilman laajempaa kuvaa ongelman käsittelystä. Heikon ja yleisen tekoälyn erona on Norvigin ja Russellin (2010) mukaan myös se, että heikko tekoäly ei pysty skaalautumaan laajoihin ja vaikeisiin ongelmatilanteisiin. Russellin (2020) mielestä (yleisen) tekoälyn mallintamisessa suurimpana haasteena on ihmisen ymmärtäminen. Ihmisten aivot ovat keskenään melko samantaisia, joten pystymme helposti eläytymään toistemme psyykkisiin ja emotionaalisiin kokemuksiin. Ihmisen ei esimerkiksi tarvitse lyödä itseään vasaralla sormeeseen voidakseen kuvitella miltä se tuntuu. Koneiden sen sijaan täytyy aloittaa tyhjästä ihmisten ymmärtämiseen tarvittavan vastaavan tiedon kartuttaminen. Vaikka koneilla olisi pääsy kaikkeen ihmisen psykologiseen ja neurologiseen tietoon, ihmisen tasoisen tai ihmisen ylittävän ymmärryksen muodostaminen ihmisestä tulee vaatimaan aikaa. (Russell, 2020.)

Shushan ja Xiaodi (2017) olivat kuusi vuotta sitten sitä mieltä, että yleinen tekoäly on vain kuvitelma Hollywoodin sci-fi -elokuvissa ja siihen uskovien mielessä. Heidän näkemyksensä mukaan toistaiseksi ei ole onnistuttu kehittämään yhtään konetta, joka todella ymmärtäisi ihmisen kieltä tai osaisi ajatella luovasti ja tuottaa omaa kulttuurista sisältöä. (Shushan & Xiaodi, 2017.) Ford (2018) selvitti vuonna 2018 valikoidulta joukolta tekoälytutkijoita, mikä olisi heidän näkemyksensä mukaan ajankohta, jolloin ainakin 50 % todennäköisyydellä olisi onnistuttu saavuttamaan yleinen tekoäly tai ihmisen tasoisen tekoäly. Fordin haastatteluissa vastausten keskiarvo oli vuosi 2099, mutta hän huomauttaa, että muissa laajemmissa tutkimuksissa tulokset ovat vaihdelleet vuosien 2040 ja 2050 välillä. Vaikka Fordin vastaajajoukko oli pieni, se edusti Fordin mukaan tekoälytutkimuksen eliittiä ja on siten kiinnostava. Käytännössä siis yleisen tekoälyn voidaan ajatella edelleen siintävän 50 tai jopa 100 vuoden päässä. (Ford, 2018.) Toisaalta Jiang ym. (2022) muistuttavat, että toistuvien iteraatioiden kautta eli parhaita lopputuloksia eteenpäin kehittämällä on mahdollista, että tekoäly ylittää ihmisen kyvyt kaikkien näkökulmien osalta. Tällaisen kehityskulun myötä voisi syntyä supertekoäly (*artificial super intelligence, ASI*) joka kykenee oppimaan kokemuksesta, sopeutuu uusiin tilanteisiin, käsittelee abstrakteja käsitteitä ja hyödyntää tietojansa ympäristön käsittelyyn. Supertekoäly kykenisi loogiseen päättelyyn, ajatteluun, strategiaan, taitojen oppimiseen ja omien tietojensa arviointiin ihmistä paremmin. Siitä, milloin tällainen tekoäly saavutetaan tai saavutetaanko koskaan, ei ole yhtenäistä tulkintaa. (Jiang ym., 2022.) Itse jaan edellä kuvattujen tutkimustulosten mukaisen näkökulman, että yleisen tekoälyn saavuttamiseen on edelleen matkaa. Pohdin tätä asiaa tarkemmin vielä gradun lopussa suhteessa tutkimustuloksiini.

2.3 Ongelmanratkaisu monimutkaisessa ja muuttuvassa ympäristössä

Russell (2020) kuvaa älykkyyden kytkeytyneen yhteen havainnoin, päättelyn ja onnistuneen toiminnan kanssa jo antiikin kreikan filosofiasta lähtien. Länsimaissa rationaalisen ajattelun määritelmä on pohjautunut Aristoteleen mukaiseen näkemykseen, että loogista päättelyä seuraava toimintojen sarja johtaa parhaaseen haluttuun lopputulokseen. Tästä näkemyksestä jää Russellin mielestä kuitenkin huomioimatta epävarmuuden vaikutukset. Todellisessa maailmassa vain harva toiminta tai toimintojen sarja toteutuu ajatusten mukaisesti, sillä todellisuus sisältää useita epävarmuustekijöitä, jotka vaikuttavat tilanteeseen. (Russell, 2020.) Maailma ja ihmiset ovat monimutkaisia ja tilanteet jatkuvasti muuttuvia, mikä tekee myös ongelmanratkaisusta haastavaa.

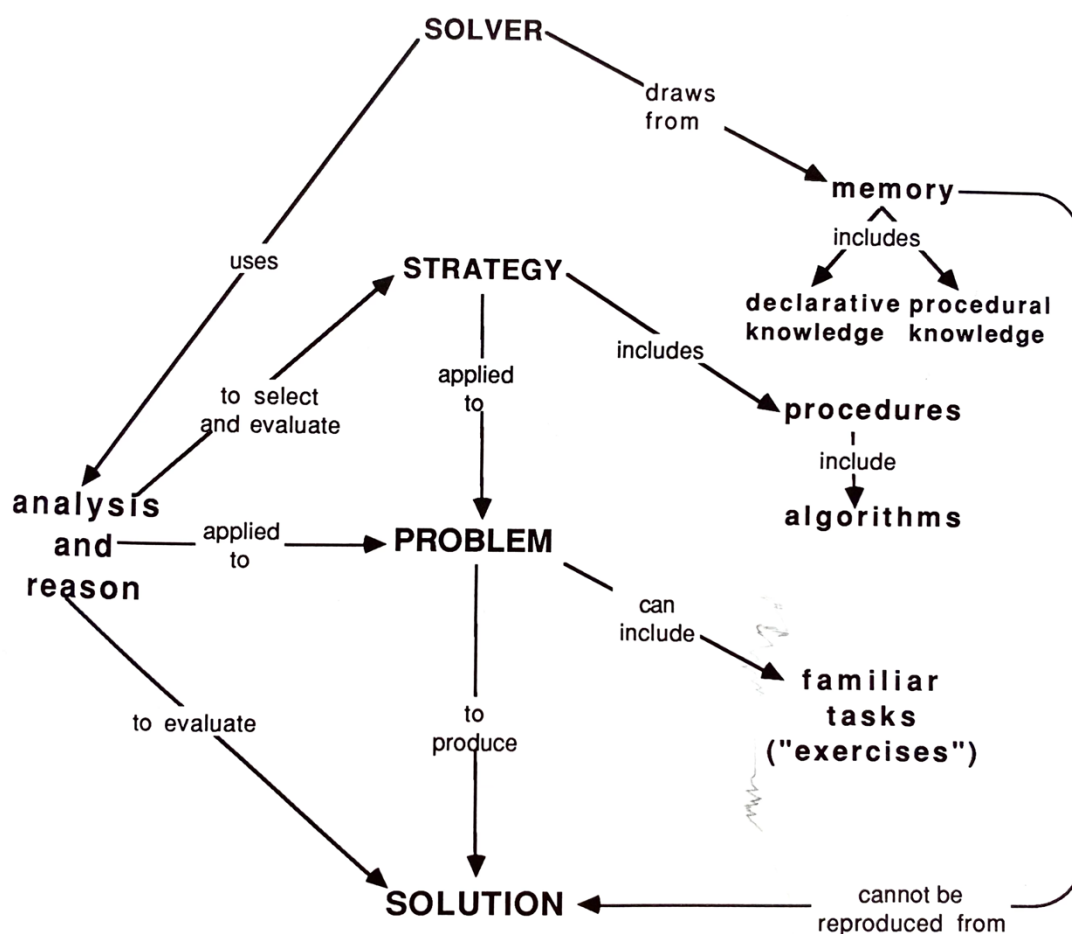
Päivittäiseen elämään sisältyy useita erilaisia ongelmanratkaisutilanteita: mitä laittaa aamulla päälle, miten mennä harrastuksiin, miten ratkaista viimeisin työtehtävä jne. Suurin osa arjessa tehtävistä päätöksistä liittyy jotenkin pieniin tai isoihin vastaan tuleviin ongelmiin, jotka pitää ratkaista tavalla tai toisella (Wang & Chiew, 2010). Käytännössä ongelmanratkaisussa on kyse siis siitä, että nykytilan ja tavoitetilan välillä on jonkinlainen este ja sen selättämiseksi ei ole heti löydettävissä ilmiselvää ratkaisua (Goldstein, 2015). Jonassenin (2000) mielestä ongelmanratkaisun kannalta on oleellista myös se, että joku havaitsee ongelman eli tuntemattoman ja kokee sen ratkaisemisen merkitykselliseksi: ongelman selvittäminen tuottaa sosiaalista, kulttuurista tai henkistä arvoa. Jos kukaan ei havaitse tuntematonta tai koe siinä mitään selvittämisen tarvetta, ongelmaa ei varsinaisesti ole. (Jonassen, 2000.) Smithin (1991) mielestä kirjallisuudesta puuttuu edelleen yhtenäinen käsitys siitä, miten ongelma määritellään. Mikään määritelmä ei ole saanut laajaa hyväksyntää tutkimuskentällä, vaan määritelmiä on käytössä lähes yhtä paljon kuin on tutkijoitakin. Smith itse määrittelee ongelman kuvion 4 mukaisesti eli ongelma on mikä tahansa tehtävä, joka edellyttää analysointia ja päättelyä tavoitteeseen tai ratkaisuun pääsemiseksi. Analysoinnin ja päättelyn tulee perustua tehtäväalueen ymmärrykseen ja ongelmaa ei voida ratkaista ainoastaan muistelemalla, tunnistamalla tai jäljittelemällä. Ongelmanratkaisu on siis prosessi, jossa kehitetään hyväksyttävä ratkaisu ongelmaan. (Smith, 1991.)

Wang ja Chiew (2010) kuvaavat ongelman muodostuvan tosiasioista, tavoitteista ja toiminnasta. Tosiasiat ovat tietoa, joka on saatavilla osana ongelmaa, tavoitteita voidaan pitää ongelman toivottuna päätepisteenä, ja toiminnat ovat ratkaisun saavuttamiseksi tehtäviä mahdollisia tekoja. (Wang & Chiew, 2010). Smith (1991) huomauttaa, että mikä tahansa tehtävä ei ole ongelma, vaan ainoastaan tehtäviä, jotka ovat riittävän monimutkaisia, voidaan pitää todellisina ongelmina. Keittiöpöydän äärestä sohvalle siirtyminen ei edellytä ongelmanratkaisua, ellei tilanteessa ole jotain poikkeavaa, kuten esimerkiksi liikkumisen rajoitteita tai sohvan sijainti epätavallisessa paikassa.

Condellin ym. (2010) mielestä ongelmien osalta on kaksi kategoriaa: tarkkaan määritellyt ja epäselvät. Tarkkaan määritellyssä ongelmassa ongelmanratkaisijalla on kaikki oleellinen informaation ongelman ratkaisemiseksi, kun taas

epäselvässä ongelmassa on vain vähän tai ei lainkaan informaatiota. (Condell ym., 2010.) Myös Wang ja Chiew (2010) ovat sitä mieltä, että ongelmat voidaan jakaa kahteen kategoriaan, mutta heidän jakonsa perustuu ongelman laajuuteen ja siihen liittyviin ratkaisuvaihtoehtoihin. 1) Konvergentti eli yhteen ideaan keskittyvä ongelma, jossa tavoite on selkeä, mutta miten tavoitteeseen päästään on epäselvää. 2) Divergentti eli moniin ratkaisuvaihtoehtoihin perustuva ongelma, jonka osalta ratkaisut voivat olla joko tiedossa tai epäselviä. (Wang & Chiew, 2010.) Vaikka Wang ja Chiew eivät jaottelussaan suoraan mainitse monimutkaisia ongelmia, vaihtoehto kaksi kuvastaa mielestäni monimutkaisia tai jäsentymättömiä ongelmia, koska siihen voi liittyä useita erilaisia tekijöitä.

Tässä tutkimuksessa monimutkaisella ja jäsentymättömällä ongelmallalla tarkoitetaan edellisten kuvausten mukaisesti sellaista ongelmaa, johon ei ole olemassa yhtä selkeää ratkaisua ja siihen vaikuttavat tekijät ovat epämääräisiä ja voivat muuttua. Tällainen ongelma oli esimerkiksi vuonna 2020 levinnyt Covid19 -pandemia. Jonassenin (2000) mielestä kaikkein monimutkaisimmat ongelmat ovat muuttuvia eli tehtäväympäristö ja siihen vaikuttavat tekijät muuttuvat ajan myötä. Monimutkaisuus liittyy siihen, kuinka monia erilaisia ja muuttuvia tekijöitä tai toimintoja ongelmaan liittyy ja miten selkeästi ne ovat pääteltävissä ongelmosta. (Jonassen, 2000.)



Kuvio 4. Ongelman visuaalinen kuvaus Smith (1991) mukaan

2.4 Ihminen ongelmanratkaisijana

Wang ja Chiew (2010) määrittelevät ongelmanratkaisun olevan muuntuva ja monimuotoinen kognitiivinen prosessi, joka toimii vuorovaikutuksessa lähes kaikkien muiden mentaalisten prosessien kanssa. Myös Condell ym. (2010) ovat samoilla linjoilla sanoessaan, että erilaiset ongelmat edellyttävät aivojen eri osalueiden hyödyntämistä, joten elinaikanamme käytämme todennäköisesti kaikkia suurimpia aivojen alueita ongelmien ratkomiseen. Ongelmanratkaisua koskeva tutkimus on siis keskittynyt määrittelemään psyykkisiä prosesseja, joita ongelmanratkaisutilanteessa ilmenee sekä sitä miten ongelmanratkaisusta voitaisiin tehdä helpompaa (Goldstein, 2015).

Ensimmäiset tieteelliset maininnat ongelmanratkaisusta ovat Holthin (2008) mukaan vuodelta 1900, mutta vasta 1920-luvulta eteenpäin tutkimusten määrä alkoi varsinaisesti lisääntyä. Vuonna 1966 ongelmanratkaisua koskevan tutkimuksen osalta tunnistettiin neljä erilaista alaryhmää: (1) Gestalt/kognitiivinen näkökulma, (2) oppimisen näkökulma (3) tietotekninen/tietojenkäsittelyllinen näkökulma ja (4) psykometrinen/komponenttianalyysinen näkökulma. (Holth, 2008.) Myös Wang ja Chiew (2010) toteavat, että ongelmanratkaisua on tutkittu psykologiassa, kognitiotieteessä sekä laskennallisessa älykkyydessä useista erilaisista näkökulmista, kuten

- Suorat faktat – oikean ratkaisuvaihtoehdon löytäminen tiedossa olevien ratkaisujen pohjalta
- Heuristiikat – peukalosääntöjen tai parhaan mahdollisen ratkaisun hyödyntäminen
- Analogiat – uuden ongelman pelkistäminen sellaiseen muotoon, että tiedossa olevia ratkaisuja on mahdollista hyödyntää
- Kukkulan kapuaminen – valitaan ratkaisuja, jotka vievät pikkuhiljaa lähemmäs päämäärää
- Algoritminen päättely – sovelletaan tiedossa olevaa ja tarkkaan määriteltyä ratkaisua ongelmaan
- Perusteellinen haku – systemaattisen haun avulla selvitetään kaikki mahdolliset vaihtoehdot
- Hajota ja hallitse – ongelman ratkaiseminen jakamalla se pienempien osaongelmien joukkoon
- Analyysi ja synteesi – ongelman pienentäminen tiedossa olevaan kategoriaan ja sitten sopivan ratkaisun etsiminen

Minkään edellä mainittujen näkökulmien hyödyntäminen ei kuitenkaan takaa Wangin ja Chiewin mukaan ratkaisua ongelmalle, varsinkaan jos ratkaisu ei ole ongelmanratkaisijan ongelma-avaruudessa. (Wang & Chiew, 2010.) Käsite ongelma-avaruudesta aukeaa tarkemmin seuraavissa kappaleissa, joissa esitellään tarkemmin muutamia keskeisimpiä kognitiotieteellisiä teorioita ongelmanratkaisusta ja miten ne tulkitsevat ongelmanratkaisun tapahtuvan ihmisen mielessä.

2.4.1 Gestalt-psykologit ja älynväläys

Gestalt-psykologit olivat ensimmäisten joukossa toteuttamassa ongelmanratkaisua koskevaa tutkimusta 1920-luvulla (Condell ym., 2010). Gestalt-psykologian mukaisen lähestymistavan ajatuksena on, että onnistuminen ongelmanratkaisussa on riippuvaista siitä, minkälainen representaatio⁴ asiasta on mielessä (Goldstein, 2015). Gestalt-psykologinen lähestymistapa ongelmanratkaisuun keskittyi erityisesti tarkastelemaan ongelman alkuperäisen representaation ja ongelman ratkaisemisen yhteydessä muodostuvan uuden representaation prosessia, jota he kutsuivat representaation uudelleen järjestelyksi (Condell ym., 2010; Goldstein, 2015). Gestalt-psykologit uskoivat, että ihmiset kokivat ongelmia ratkoessaan älynväläyksen, koska lopullinen ratkaisu vaikutti tulevan heidän mieleensä yhtäkkiä. Gestalt-psykologien näkökulmat ovat myöhemmissä tutkimuksissa saaneet sekä kannatusta että kritiikkiä. (Goldstein, 2015.)

2.4.2 Hakuun perustuva ongelmanratkaisu

Alan Newell ja Herbert Simon lähestyivät ongelmanratkaisua tietojenkäsittelyllisestä näkökulmasta. He kehittivät loogisteoreettisen ohjelman (*General Problem Solver* eli GPS), joka oli suunniteltu simuloimaan ihmisen ongelmanratkaisua (Goldstein, 2015). Jonassen (2000) kuvaa, että GPS määrittelee ongelmanratkaisuun yhdistyvän kaksi ajattelun prosessia, (a) ymmärtäminen ja (b) hakuprosessi. Käytännössä ideana on tutkia ja arvioida erilaisten vaihtoehtojen seurauksia ja valita siltä pohjalta paras ratkaisu. Goldstein (2015) kuvaa Newellin ja Simonin ongelmanratkaisuteorian sisältävän sarjan valintoja, joista jokainen muodostaa oman vaiheensa. Ongelmanratkaisun eri vaiheet, alkuvaiheesta välivaiheiden kautta tavoitettiin, muodostavat ongelma-avaruuden (*problem space*). Löytäkseen ratkaisun ongelmaan, ongelmanratkaisija ikään kuin suorittaa haun ongelma-avaruudessa. (Goldstein, 2015.) Ongelmanratkaisu voidaan siis nähdä mielensisäisenä hakuprosessina, jossa muistista yritetään löytää jonkinlaisia yhteyksiä vaihtoehtoisten ratkaisujen ja lopputulosten välillä (Kolak ym., 2006). Langley ja Rogers (2005) kuvaavat samaa asiaa tiedon järjestelemisenä ratkaisun löytämiseksi.

Newellin ja Simonin teoria käynnisti Goldsteinin (2015) mukaan tutkimussuuntauksen, joka kuvaa ongelmanratkaisun prosessina, johon sisältyy tiedon hakua sekä ongelman muotoilun että ratkaisun etsinnän osalta. Tutkimukset ovat kuitenkin osoittaneet, että ongelmanratkaisuun liittyy muutakin kuin GPS:n keskiössä olevan ongelma-avaruuden määrittely. (Goldstein, 2015.)

Jonassen (2000) tosin on sitä mieltä, että mentaalinen tulkinta ongelma-avaruudesta on ongelmanratkaisussa oleellista ja ongelmanratkaisu edellyttää jonkinlaista ongelma-avaruuden toimintokohtaista manipulointia. Vaikka tietojenkäsittelyllinen näkökulma ei tarjoa kaiken kattavaa teoriaa ongelmanratkaisusta, on siinä Gestalt-psykologisen teorian tapaan elementtejä, joita on sovellettu

⁴ Representaatiot ovat yksilöllisiä mielen sisäisiä mielikuvia, malleja tai käsityksiä asioista.

useissa myöhemmissä ongelmanratkaisua koskevissa teorioissa tai näkökulmissa.

2.4.3 Analoginen ongelmanratkaisu

Eräs yleinen ongelmanratkaisuteoria on Condell ym. (2010) mukaan analogioiden hyödyntäminen ongelmanratkaisussa Thagard (2005) määrittelee analogioiden olevan ihmisen aiempiin onnistuneisiin ja epäonnistuneisiin kokemuksiin pohjautuvia ratkaisu- tai toimintamalleja. Käytännössä analogioita hyödynnetään jonkin samankaltaisen ongelmatilanteen ratkaisemiseen (Condell ym., 2010). Goldstein (2015) kuvaa analogista ongelmanratkaisua esiintyvän erityisesti silloin, kun aiemmin ratkaistua ongelmaa tai tarinaa voidaan hyödyntää tietolähteenä uuden ongelman ratkaisemiseen. Reeves ja Weisberg (1994) kuvailevat analogisen siirron olevan tärkeässä roolissa analogisessa ongelmanratkaisussa. Analoginen siirto rakentuu neljän vaiheen pohjalta a) Pohja- ja kohdeanalogioiden tunnistaminen b) Pohja-analogian mieleen palauttaminen c) Pohja-analogian soveltaminen tai kartoitus suhteessa kohdeanalogiaan d) Ratkaisuperiaatteiden yhteensovittaminen kohdeongelmaan. Viimeinen vaihe ei tosin ole tutkimuksissa yhtä laajasti tunnistettu, kuin kolme ensimmäistä. Analogioiden siirtoprosessiin on erilaisia teoreettisia lähestymisiä, kuten Gentnerin *structure-mapping model* ja Holyoakin ja kumppaneiden *pragmatic schema theory* skeemateoria. Ajatuksena näissä on, että analogioiden siirtoa uuteen ongelmaan voidaan helpottaa käsitteellisen kartoituksen tai skeemojen avulla. (Reeves ja Weisberg, 1994.) Gick ja Holyoak (1983) kuvaavat, että analogisen ajattelun keskiössä on tiedon siirtäminen yhdestä tilanteesta toiseen kartoitukseen (mapping) perustuvan prosessin kautta eli yhteyksien löytäminen eri tarkastelukohteiden välillä.

Vaikka ihmisten on usein vaikea tutkimuksellisissa koetilanteissa hyödyntää analogioita ongelmanratkaisuun, käytännön kenttätutkimuksissa on tullut ilmi, että tosielämän tilanteissa analogioita hyödynnetään usein. (Goldstein, 2015.) Myös Condell ym. (2010) mainitsevat analogioiden hyödyntämiseen liittyvän haasteita erityisesti silloin, kun oikea ratkaisu edellyttää pitkää toimintojen sarjaa. Analogiat ovat toimivimpia tilanteissa, joissa ratkaisu on helppo muistaa. (Condell ym., 2010.) Thagardin (2005) mielestä analogiat eivät myöskään sovi täysin uusiin tilanteisiin, koska niiden avulla saattaa päätyä käyttämään sellaista analogiaa, joka ei ole välttämättä paras mahdollinen kyseisen ongelman ratkaisemiseksi.

2.4.4 Yksilölliset erot ongelmanratkaisukyvyssä

Vaikka ongelmanratkaisu on arkipäivää kaikille ihmisille, kaikki eivät omaa samanlaisia ongelmanratkaisutaitoja; jollekulle haastava ongelma ei välttämättä ole vaikea toiselle (Wang & Chiew, 2010). Yleisen käsityksen mukaan osa ihmisistä on parempia ongelmanratkaisijoita, koska he käyttävät tehokkaampia ongelmanratkaisustrategioita (Jonassen, 2000). Condellin ym. (2010) mukaan ekspertit luottavat ongelmanratkaisussa aiemmista kokemuksista kerryttämäänsä tietoon. Smith (1991) tosin huomauttaa, että ekspertejä ja noviiseja koskevissa tutkimuksissa on säännöllisesti pystytty tunnistamaan myös noviiseja, joiden

ongelmanratkaisutekniikat ovat hyvin vastaavia kuin eksperteillä. Thagard (2005) sanoo, että ratkaistessaan ongelmaa, ihmiset yleensä oppivat kokemuksesta ja pystyvät siten ratkaisemaan vastaavia ongelmia helpommin jatkossa.

Wangin ja Chiewin (2010) mukaan ongelmanratkaisukyvyyn yksilölliset erot voidaan selittää empiirisesti asiantuntemuksen, tietopääoman ja ongelmaratkaisutaidon pohjalta. Vaikka kaikille annettaisiin sama ongelma, ongelmanratkaisuvaihtoehdot olisivat erilaiset henkilöiden kokemuksesta, tiedoista, taidoista ja ongelmanratkaisustrategioista johtuen. Jokaista ongelmaa kohden on mahdollinen ongelma-avaruus, joka pitää sisällään ongelmanratkaisijan saavutettavissa olevat mahdolliset vaihtoehdot ja ratkaisut. Käytännössä tämä tarkoittaa sitä, että ongelmanratkaisijasta riippuen, kaikki ratkaisuvaihtoehdot eivät aina ole mahdollisia. Ratkaisu tiettyyn ongelmaan ei välttämättä ole mukana ongelma-avaruudessa, koska ongelma voi olla jäsentymätön, tavoitetilä epäselvä tai sopivaa menetelmää ei ole saatavilla. (Wang & Chiew, 2010.) Samoilla linjoilla on myös Jonassen (2000), joka uskoo ongelmanratkaisutaidon muodostuvan ongelman luonteesta, minkälaisen representaation ratkaisija on muodostanut ongelmasta ja yksilön henkilökohtaisista ominaisuuksista.

2.4.5 Eri teorioita yhdistävät mallit

Jonassen (2000) sanoo, että ongelmanratkaisumalleilla on tapana käsitellä kaikkia ongelmia samalla tavalla, koska ne pyrkivät ilmaisemaan yleistettävissä olevaa ongelmanratkaisuprosessia. Jonassenin mielestä ongelmanratkaisu ei kuitenkaan ole samanlaisena toistuvaa toimintaa, koska ongelmatkaan eivät ole sisällöltään tai muodoltaan yhteneväisiä. (Jonassen, 2000.) Myös Smith (1991) tuo esille, että joidenkin tieteenalojen tutkimuksissa on korostettu ongelmanratkaisun tilannesidonnaisuutta ja esitetty että eri aloilla ongelmanratkaisusta on löydetty enemmän eroavaisuuksia kuin yhteneväisyyksiä.

Ongelmanratkaisun tilannesidonnaisuudesta huolimatta Smith ja kumppanit yrittivät muodostaa yhtenäisen teorian ongelmanratkaisusta, tosin huomaten nopeasti, ettei se tulisi onnistumaan. Smith laati kuitenkin synteysin ongelmanratkaisuprosessista toivoen tutkimuskentän siten pääsevän lähemmäs yhtenäistä teoriaa. (Smith, 1991.) Myös Wang ja Chiew (2010) ovat laatineet oman ehdotuksena ongelmanratkaisun kognitiivisesta prosessista (kuvio 5). Esittelen alla lyhyesti molemmat prosessikuvaukset, koska mielestäni niiden kautta välittyy hyvin, miten aiempia ongelmanratkaisua koskevia teorioita on yhdistetty nykyiseen käsitykseen ongelmanratkaisusta. Molemmat teoriat kuvaavat mielestäni hyvin myös tiettyjä ihmisen ongelmanratkaisun kannalta keskeisiä piirteitä, joita tekoälyllä on ollut ja tulee uskoakseni myös jatkossa olemaan haastava saavuttaa.

Smith (1991) kuvaa hyvien ongelmanratkaisijoiden toiminnan keskeisiä piirteitä seuraavasti:

1. Ongelmanratkaisija luo sisäisen ongelma-avaruuden eli omakohtaisen ymmärryksen ongelmasta. Tähän sisältyy representaatio ongelmasta ratkaisijan tietämyksen kannalta sopivassa muodossa, sekä ongelman yksityiskohdat sillä tavalla, että seuraavien vaiheiden toteuttaminen on mahdollisimman yksinkertaista.

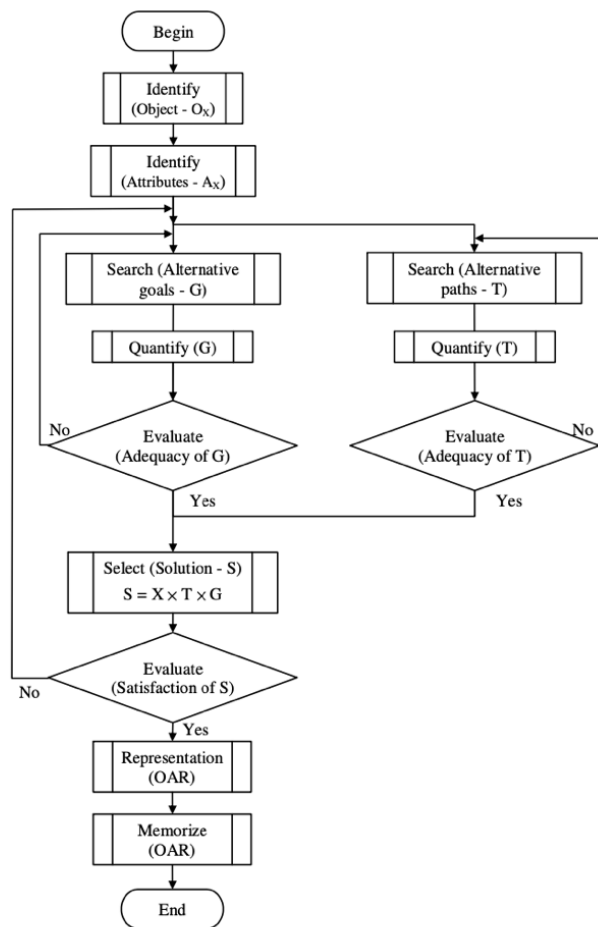
2. Riippuen ongelman monimutkaisuudesta, menestyvä ongelmanratkaisija suunnittelee peruslinjat tai yleisen strategisen lähestymistavan ratkaisuprosessin etenemisestä. (Tämä voi jäädä pois helpossa ongelmanratkaisutilanteessa)
3. Seuraavaksi hyvä ongelmanratkaisija soveltaa tehtävään sopivaa menettelytapaa ongelman ratkaisemiseen. Tässä yhteydessä ongelmanratkaisija käyttää kahdenlaisia heuristiikkoja:⁵ yleisiä heuristiikkoja tai alakohtaisia heuristiikkoja ja algoritmeja. Yleiset tai heikot heuristiikat ovat laajasti käyttökelpoisia, mutta tyypillisesti tehottomia sopivan ratkaisun saavuttamisessa nopeasti. Alakohtaiset heuristiikat ja algoritmit ovat yleensä sopivia ainoastaan rajatun alan ongelmiin. Ne ovat toimiva nopean ja paikkansapitävän ratkaisun löytämisessä, mutta ne edellyttävät pätevää sisällön ymmärrystä alalta.
4. Pätevät ongelmanratkaisijat pystyvät tarvittaessa toteuttamaan monivaiheisia menettelyjä, pitäen samalla mielessä edelliset vaiheet.
5. Viimeisenä vaiheena parhaimmat ongelmanratkojat tekevät vielä ratkaisun jälkiarvioinnin, joka voidaan toteuttaa eri tavoin: kvalitatiivinen arviointi, paikkansapitävyyden tarkistaminen, aikaisempiin ratkaisuihin vertaaminen.

Wang ja Chiew (2010) mukaan ongelmanratkaisun kognitiivinen prosessi voidaan kuvata viiden vaiheen kautta:

1. Ongelman määrittely eli keskeisten ominaisuuksien havainnointi ja tavoitteen tunnistaminen.
2. Tavoitteiden ja reittien haku: Aivot muodostavat rinnakkaisia hakuja löytääkseen mahdollisia reittejä ja tavoitteita ratkaisuun pääsemiseksi. Jos ongelmanratkaisijan sisäisissä tiedoissa ei ole sopivia tavoitteita tai reittejä, hyödynnetään ulkoisia tietolähteitä.
3. Mahdollisten ratkaisujen kehittäminen.
4. Soveltuvan ratkaisun valinta, joka tapahtuu mahdollisten ratkaisujen arvioinnin pohjalta. Tarvittaessa toteutetaan lisähakuja, mikäli tunnistetut ratkaisut eivät täytä ongelmanratkaisijan vaatimuksia.
5. Representaation muodostaminen ongelmanratkaisusta. Representaatio tallennetaan kokonaisuudessaan pitkäkestoiseen muistiin.

Smith (1991) sekä Wang ja Chiew (2010) esittämät ongelmanratkaisuprosessit ovat mielestäni peruseräillä hyvin samanlaiset. Molemmat lähtevät liikkeelle ongelman määrittelystä, sisältävät ratkaisujen suunnittelun ja kehittämisen sekä päättyvät ratkaisujen arviointiin. Tarkemman tarkastelun pohjalta tulkitaisin Wang ja Chiew (2010) kuvauksen olevan yleistettävämpi ja kuvaavan selkeämmin ongelmanratkaisun kognitiivista prosessia, sekään ei silti ole kattava kuvaus ongelmanratkaisuprosessista.

⁵ Heuristiikat ovat ihmisen mielensisäisiä sääntöjä tai malleja, joita hyödyntämällä ongelmanratkaisutilanteessa ei tarvitse käydä läpi kaikkia mahdollisia vaihtoehtoja (Thagard, 2005).



Kuvio 5. Ongelman kognitiivinen prosessi Wang ja Chiew (2010) mukaan

2.5 Tekoäly ongelmanratkaisijana

Kappaleessa 2.2.2 esittelin lyhyesti erilaisia tekoälymenetelmiä, joita hyödynnetään monenlaisilla sovellusalueilla. Myös tekoälyn ongelmanratkaisuun on käytössä erilaisia algoritmeja, esimerkiksi sellainen, jolle ratkaistavasta ongelmasta ei anneta muuta tietoa kuin sen määritelmä. Tällaiset algoritmit pystyvät kyllä ratkomaan minkä tahansa ratkaistavissa olevan ongelman, mutta eivät kovin tehokkaasti. Informoidut hakualgoritmit sen sijaan pärjäävät paremmin, jos niille annetaan ohjausta sopivan ratkaisun löytämiseksi. (Norvig & Russell, 2010.) Käytännössä tekoälyn ongelmanratkaisun tutkimus on jakautunut useisiin tieteenaloihin, jotka kyllä jakavat olennaiset ongelmanratkaisua koskevat näkökulmat, mutta keskittyvät toisistaan eroaviin sovelluksiin (Luger, 2005).

Johnson ym. (2022) toteavat, että valtaosalla algoritmeista on sellaiset sisäiset rakenteet, että niiden tekemät ratkaisut eivät ole tulkittavissa tai selitettävissä,

erityisesti jos algoritmia on koulutettu musta laatikko⁶ (*black box*) -menetelmällä. Ongelmanratkaisun kannalta on kuitenkin oleellista ymmärtää, miten ratkaisuun on päädytty, jotta voidaan esimerkiksi arvioida tekoälyn hyödyntämismahdollisuuksia.

Norvig ja Russell (2010) kuvaavat miten yksinkertainen toimija (agentti) voisi edetä ongelmanratkaisussa. Funktioksi muotoiltuna ongelmanratkaisu voisi edetä listauksen 1 mukaisesti. Ensin toimija muotoilee ongelman ja tavoitteen (*formulate*), sitten kartoittaa mahdollisia ongelmanratkaisun kannalta oleellisia toimintavaihtoehtoja (*search*) ja lopuksi toteuttaa toiminnot yksi kerrallaan (*execute*). Saatuaan prosessin päätökseen, se muotoilee uuden päämäärän ja aloittaa jälleen alusta. (Norvig & Russell, 2010.)

function SIMPLE-PROBLEM-SOLVING-AGENT (*percept*) returns an action persistent:

seq, an action sequence, initially empty state
state, some description of the current world state goal
goal, a goal, initially null problem
problem, a problem formulation

state ← UPDATE-STATE (*state*, *percept*)

if *seq* is empty then

goal ← FORMULATE-GOAL(*state*)

problem ← FORMULATE-PROBLEM(*state*, *goal*)

seq ← SEARCH(*problem*)

if *seq* = failure **then return** a null action

action ← FIRST(*seq*)

seq ← REST(*seq*)

return *action*

Listaus 1 Ongelmanratkaisuprosessin funktio

Hieman listauksen 1 funktiota tarkemmin kuvattuna ongelmanratkaisu voitaisiin Norvigia ja Russellia (2010) tiivistäen kuvata seuraavasti:

1. Ongelmanratkaisun ensimmäinen vaihe on tavoitteenmuodostus, joka perustuu nykytilanteeseen sekä toimijan⁷ suorituskykyyn. Tavoitteet auttavat toiminnan järjestelyssä rajoittamalla harkittavien toimintojen määrää.
2. Toisena vaiheena on ongelman muotoilu, jossa pyritään selvittämään miten kannattaa toimia nyt ja jatkossa, jotta saavutetaan tavoitetila. Toimijan on päätettävä (tai meidän on päätettävä sen puolesta) millaisia toimia ja tiloja sen kannattaa harkita.

⁶ Mustalaatikko ongelma tarkoittaa sitä, että kukaan ei tiedä miten ja miksi tekoäly on päätenyt tiettyyn lopputulokseen.

⁷ Toimija tarkoittaa algoritmia, johon tekoälyn toiminta perustuu.

3. Kolmannessa vaiheessa toimija kokeilee yhtä toimintoa satunnaisesti, koska se ei vielä tiedä tarpeeksi toimintojen mahdollisista seurauksista. Mikäli toimijalla on useita välittömiä mahdollisuuksia, joiden merkitys ei ole tiedossa, se voi valita seuraavat toimet tutkimalla ensin mahdollisia vaihtoehtoja. Osa vaihtoehtoista etenemismahdollisuuksista voidaan myös jättää pois, koska ne ovat epäoleellisia lopullisen ratkaisun näkökulmasta. Yksityiskohtien poistamista representaatiosta kutsutaan erottamiseksi (*abstraction*). Erottamisessa pyritään poistamaan niin paljon yksityiskohtia, kuin mahdollista, kuitenkin säilyttäen luotettavuus ja toimintojen helppo toteuttaminen. Ilman onnistunutta erottamista, toimija hukkuisi tosimaailman mahdollisiin vaihtoehtoihin.
4. Lopullinen ratkaisu muodostuu hakupuun perusteella. Hakualgoritmit etenevät hakupuussa alkaen lähtötilasta kohti vaihtoehtoisia haaroja, risteyskohdat vastaavat hakuavaruutta, jossa harkitaan erilaisia mahdollisia toimintavaihtoehtoja.
5. Viimeisenä vaiheena on toteuttaminen, joka tehdään sen pohjalta, minkälainen ratkaisu haun perusteella on löytynyt. (Norvig & Russell, 2010.)

Luger (2005) mukaan tekoälytutkijoiden keskuudessa kaksi perustavanlaatuisinta huolenaihetta ovat tiedon representaatiot ja haku. Tekoälyn ongelmanratkaisun yhteydessä representaatio tarkoittaa sellaisessa muodossa tallennettua tietoa, jota kone pystyy tulkitsemaan (esim. pikselit ja numerot). Representaatiokaavioihin tallennetaan ongelmatilanteen olennaiset ominaisuudet, jotta tarvittava informaatio on ongelmanratkaisutilanteen osalta saavutettavaa. (Luger, 2005.) Marcus ja Davis (2019) mielestä tekoälyllä on representaatioita ainoastaan teknisestä näkökulmasta tarkasteltuna: syötteitä ja tulosteita edustavia numerosarjoja eli vektoreita, mutta niiltä puuttuu kokonaan representaatiot laajemmasta näkökulmasta tarkasteltuna.

Haulla tarkoitetaan ongelmanratkaisutekniikkaa, jossa systemaattisesti tarkastellaan mahdollisten tilojen eli vaihtoehtojen avaruutta (Luger, 2005). Tila-avaruus määrittyy ongelmaan sisältyvän alkutilan, toimintojen ja siirtymien perusteella. Tila-avaruushaussa muodostetaan alkutilasta haarautuva vektori tai graafi, jossa solmut ovat mahdollisia tiloja ja linkit solmujen välillä ovat toimintoja. Tila-avaruushaun avulla kartoitetaan kaikki alkutilasta saavutettavissa olevat tilat. (Norvig & Russell, 2010.) Tila-avaruushaku ei kuitenkaan toimi kaikissa ongelmanratkaisutilanteissa, sillä kaikkien vaihtoehtoisten tilojen kartoittaminen olisi käytännössä mahdotonta (Luger, 2005). Myös Lake ym. (2017) toteavat, että laajan hypoteesiavaruuden osalta on mahdoton tehdä tyhjentävää hakua, jolla löytyisi toimiva ratkaisu. Seuraavassa kappaleessa perehdyn vielä tarkemmin tekoälyn ongelmanratkaisussa tunnistettuihin haasteisiin.

2.5.1 Ongelmat tekoälyn ongelmanratkaisukyvyssä

Nykyisen tekoälyn keskeisin ongelma on sen kapea-alaisuus (*narrow AI*): se toimii hyvin vain niissä tehtävissä, joihin se on ohjelmoitu. Tämä on riittävää

ongelmanratkaisukykyä tilanteissa, joissa on selkeät säännöt ja logiikka, mutta tosielämän tilanteissa tarvitaan joustavaa kykyä ratkoa uusia ongelmia ja sopeutua muuttuviin tilanteisiin. (Marcus & Davis, 2019)

Tekoälyn ongelmanratkaisukyvyssä on tunnistettu useita haasteita, joita kirjallisuudessa on luokiteltu eri tavoin. Yhdenmukaista ja kattavaa luokittelua tekoälyn ongelmista en onnistunut löytämään edes pitkän ja moniulotteisen haun pohjalta, vaan eri lähteissä ongelmia oli tunnistettu ja määritelty eri tavoin. Esimerkiksi Norvig ja Russell (2010) vertaavat ihmisen ja tekoälyn ominaisuuksia ja ovat tunnistaneet tekoälyn osalta seuraavia haasteita:

1. Kyky todennäköisyyksien päättelyyn epämääräisissä olosuhteissa
2. Osiin jakaantuneet representaatiot, jotka eivät vastaa objekteja ja suhteita kattavasti
3. Yleisten ja mukautuvien representaatiomallien kehittäminen monimutkaisiin tilanteisiin
4. Tulevan toiminnan suunnittelu, arviointi ja valinta
5. Päälekkäisten etujen asettaminen paremmuusjärjestykseen
6. Uusien representaatioiden oppiminen, etenkin kun kyseessä on abstrakti syöte

Vaikka Norvigin ja Russellin lista haasteista on asianmukainen ja perusteltu, se ei mielestäni ole yhtä selkeä, kuin Lu ym. (2018) laatima kooste viimeaikaisten tekoälymallien tyypillisistä haasteista. Heidän mukaansa viimeaikaisen tekoälyteknologian keskeiset haasteet ovat frame-ongelma, assosiaatiofunktion ongelma, symbolisten merkitysten ongelma sekä mentaalisen ja fyysisen todellisuuden ymmärtämisen ongelma. Näistä frame-ongelma ja symbolisten merkitysten ongelma ovat muutenkin laajasti tunnistettuja tekoälyn haasteita. Lu ym. (2018) luokittelu onkin siten hyvä pohja myös tähän tutkimukseen, vaikka he ovatkin artikkelissaan määritelleet ongelmat hyvin ylimalkaisesti. Esittelen mainittuja ongelmia tarkemmin seuraavissa kappaleissa, mutta jätän pois mentaalisen ja fyysisen ongelman ja nostan sen tilalle tiedon hankinnan ongelman (*knowledge acquisition problem*). Koen, että Lu ym. (2018) eivät olleet määritelleet mentaalista ja fyysistä ongelmaa riittävällä tarkkuudella, eikä se noussut muissakaan lähteissä selkeästi esille, kun taas tiedon hankinnan ongelmaa on muualla kirjallisuudessa käsitelty paljon.

Frame-ongelma

Haastetta erottaa tilanteen kannalta oleellisia asioita kutsutaan frame-ongelmaksi (mm. Dennett, 1984). Frame-ongelmassa on kyse haasteesta tunnistaa, mitkä ominaisuudet avoimessa tehtäväympäristössä muuttuvat tietyn toiminnan jälkeen (Dietrich & Fields, 2020; Russell, 2020). Käytännön elämässä ilmiöt ja tilanteet sisältävät valtavan määrän muuttujia. Kaikkien muuttujien ja niiden mahdollisten tilojen ennakointi johtaisi tietokannan ylikuormittumiseen ja rajattomaan tiedonhaku aikaan (Lu ym., 2018). Dietrich ja Fields (2020) mielestä frame-ongelma koskee myös ihmistä: toimintojen seurausten ja vaikutusten ennakointi on haastavaa. Mahdolliset frame-ongelman ratkaisut perustuvat

heuristiikkaan ja ovat siten alttiita ennakoimattomille virheille. (Dietrich & Fields, 2020.) Toisaalta Kolak ym. (2006) mukaan useat filosofit ja kognitiotieteilijät uskovat, että yksi kaikkein perustavanlaatuisista ja erityisimmistä ihmisen kyvyistä on kykyimme suunnitella ja kuvitella mielessä toimintamme seurauksia.

Russellin (2020) mukaan frame-ongelmaan liittyy myös pätevyysongelma, joka tarkoittaa sitä, että jokaiselle säännölle on olemassa lukemattomia poikkeuksia. Jotta voitaisiin ennustaa, miten yksinkertaisen tilanne muuttuisi pientenkin muutosten jälkeen, tarvittaisiin järjestelmään loputon määrä sääntöjä, jotka olisi päteväitä loputtomien tavoin. (Russell, 2020.)

Assosiaatiofunktion ongelma

Nykyiset tekoälyjärjestelmät ovat riippuvaisia suurista tietoaaineistoista ja ne pystyvät tuottamaan tuloksia ainoastaan numeerisesti. Niiltä puuttuu kuitenkin ihmisille tyypillinen kyky assosiaatioiden eli asioiden välisten suhteiden ymmärtämiseen. (Lu ym., 2018.) Lucci ym. (2015) mukaan assosiaatioita voisi helposti pitää vain elämäkokemukseen perustuvina hyvinä tai huonoina muistoina, mutta todellisuudessa ne kuvaavat ihmisen ainulaatuista kykyä yhdistää vähäistä tiedoista teorioita tai ratkaisuja.

Käsittääkseni Marcus ja Davis (2019) puhuvat hieman vastaavasta ongelmasta, sanoessaan, että tekoälyn monimutkaisen ongelmanratkaisun keskeisimpiä puutteita on kyky ymmärtää propositioita⁸. Käytännössä tekoälyn representaatiot ovat ainoastaan suoraviivaista dataa ja niistä puuttuu syvällisempi ymmärrys asioiden välisistä yhteyksistä ja abstrakteista merkityksistä. Ilman ihmisen kaltaista kykyä yleistämiseen ja asioiden abstraktien representaatioiden ymmärtämiseen, tekoäly ei kykene monimutkaiseen ongelmanratkaisuun. (Marcus & Davis, 2019.)

Symbolisten merkitysten ongelma

Symbolit ovat objektien eli paikallisten, rajoitettujen aineellisten järjestelmien merkityksiä, joihin viitataan luonnollisen kielen käsitteillä ja mentaalilla representaatioilla. Symbolin viittauksen määrittäminen edellyttää siten myös kohteen tai kohteiden tunnistamista. (Dietrich & Fields, 2020.) Lu ym., (2018) kertovat symbolisten merkitysten ymmärtämisestä esimerkkinä ajatuksen hevosesta rai-doilla, jonka ihminen tulkitsee vaivatta sepraksi, mutta koneella on vaikeuksia ymmärtää vastaavia merkitysten välisiä yhteyksiä. Harnad (1990) määrittelee symbolisten merkitysten ongelmassa olevan kyse ns. kehävaikutuksesta: symbolien merkityksiä ei ole mahdollista luoda muuten, kuin määrittelemällä ne jollakolla merkityksettömiä symboleita. Haasteena on siis luoda formaaliin symbolijärjestelmään liittyvät semanttiset tulkinnat: ihmiselle symboleilla on sellaisia

⁸ Propositioiden olemassaolosta ei ole konsensusta, joten niille ei ole yksiselitteistä määritelmääkään. Propositioiden ajatellaan olevan ajatusten osia, joihin sisältyy ymmärrys siitä, onko jokin väite tai asia tosi tai epätosi esimerkiksi "vesi on märkää". Monimutkaisissa propositioissa voi olla kyse samaan lauseeseen sisältyvistä useista väitteistä tai abstrakteista asioista esimerkiksi "kauneus on katsojan silmässä".

luontaisia ominaisuuksia, joiden määrittely tekoälylle on vaikeaa. Tähän liittyy myös Searlen kiinalaisen huoneen argumentti (ks. alaviite 3). (Harnad, 1990.)

Lake ym. (2017) nostavat yhdeksi ihmisen vahvuudeksi tekoälyyn verrattuna kyvyn oppia konsepteja. Ihminen oppii pienemmästä aineistosta ja kykenee tekemään monipuolisempia ja joustavampia yleistyksiä. Yksinkertaistenkin visuaalisten käsitteiden osalta ihmiset ovat parempia ja kehittyneempiä oppijoita, kuin merkkien tunnistukseen kehitetyt algoritmit. Lake ym. mukaan ihmisen päättelykyvyn vahvuus ja joustavuus johtuu representaatioiden olemuksesta. (Lake ym., 2017.)

Tiedon hankinnan ongelma

Yksi aikamme keskeisistä ongelmista on Lucci ym. (2015) mielestä informaation muuttaminen sellaiseen muotoon, että sitä voidaan hyödyntää älykkäässä päätöksenteossa. Tieto (*knowledge*) on korkeamman tason tapa käsitellä informaatiota monimutkaisen päätöksenteon ja ymmärryksen tuottamiseksi. Tiedon osia ovat data, faktat ja informaatio. (Lucci ym., 2015) Keskeisiä tiedon muotoja ovat tilannesidonnainen tieto, eksperttiys, maalaisjärki, meta-tiedot, semantiikka, tilapäiset tiedot ja käsitteelliset tiedot (ontologia) jne. (Breuker, 2013)

Datan räjähdysmäisen kasvun myötä oleellisen tiedon tunnistaminen kaiken datan keskeltä on haastava tehtävä. (Miao ym., 2023). Koneoppiminen ja eksperttijärjestelmät ovat osaltaan pyrkineet vastaamaan tiedon hankinnan ongelmaan. Ne ovatkin onnistuneesti pystyneet tunnistamaan datamalleja valtavista aineistoista, mutta seurauksena on usein ns. mustalaatikko tilanne (ks. alaviite 6). (Confalonieri ym., 2021.) Käytännössä tekoälyn kouluttajan tulee siis tunnistaa saatavilla oleva ja tarpeellinen tieto, ennen kuin koneoppiminen voidaan aloittaa.

Ihmisen kyky uuden tiedon hankkimiseen on siis toistaiseksi ylivoimainen: ihminen kykenee oppimaan hyvin vähäisellä harjoittelulla tai jopa ilman harjoittelua uusille malleille ja sitten hyödyntämään niitä mielivaltaisesti uusissa tilanteissa ja tavoitteissa. Ihmisen kyky tehokkaaseen oppimiseen ilmenee hyvin AlphaGon tapauksesta: ennen Lee Sedolin voittamista, se pelasi 30 miljoonaa harjoituspelejä itseään vastaan, kun Lee Sedolin voidaan arvioida pelanneen vain noin 50 000 peliä. (Lake ym., 2017.)

Ongelmanratkaisun, päätöksenteon ja toiminnan kannalta järjestelmään tulisi pystyä luomaan toimiva rakenne tiedon tallentamista, nopeaa saatavuutta ja laskennallista käsittelyä varten (Lucci ym., 2015).

Edellisissä kappaleissa esiteltyjen ongelmien jälkeen voidaan yhteenvetona todeta Norvigia ja Russellia (2010) mukaillen, että tekoälykehityksessä on edelleen vaativana tehtävänä määrittellä monimutkaisiin tilanteisiin sopivia yleisiä ja mukautuvia representaatiomalleja. Myös Goertzel ja Wang (2007) muistuttavat, että yleisen tekoälynkään ei tarvitse kyetä ratkomaan mitä tahansa ongelmia, missä tahansa tilanteissa, vaan enemmänkin omata potentiaali kehittyä minkä tahansa ongelman ratkaisijaksi. Eihän ihminenäkään kykene hallitsemaan kaikkia aloja tasaaisesti, vaan oleellista on kyky oppia ja kehittyä uuden alan osaajaksi. (Legg & Hutter, 2007). Tekoälyn ajankohtaiset ongelmat myös muuttuvat ajan kanssa. Esimerkiksi Norvig ja Russell (2010) ovat kirjassaan esitelleet lupaavia

kehitysnäkymiä aiemmin mainittujen haasteiden ratkaisemiseksi: suodatusalgoritmit (*filtering algorithms*), todennäköisyyksien ja ensimmäisen lukukerran logiikka, päättelyn ja aggressiivisen koneoppimisen yhdistävä algoritmi, hierarkiset representaatiot ja syvän käsityksen verkostot (*deep belief networks*). Myös Marcus ja Davis (2019) esittävät omia ratkaisuehdotuksia tekoälyn kehittämiseksi: representaatioiden luominen ihmisen tiedon ydinrakenteista, laajenemiskykyinen tietoarkkitehtuuri, ylhäältä alas ja alhaalta ylös pohjautuva vahavan päättelyn tekniikka ja ihmisen kaltainen olemassa olevan päälle tapahtuva oppiminen. Mikään näistä ei kuitenkaan yksin ratkaise tekoälyn ongelmanratkaisun haasteita, kuten Marcus ja Davis (2019, s.155) sanovat:

It is doubtful that any single or simple approach will suffice, in part because common sense itself is so diverse; no single technique is going to solve what the field has struggled for years.

3 TUTKIMUKSEN TOTEUTUS

3.1 Systemaattinen kirjallisuuskatsaus

Okoli ja Schabram (2010) jakavat kirjallisuuskatsaukset kolmeen ryhmään. Ensimmäiseen ryhmään he sisällyttävät ”teoreettisen taustan”, jonka tarkoituksena on luoda teoreettinen perusta ja tausta tutkimuskysymykselle. Tämä on Okolin ja Schabramin mukaan yleisin kirjallisuuskatsauksen tyyppi. Toisena ryhmänä he mainitsevat graduun sisältyvän kirjallisuuskatsauksen ja kolmanneksi ryhmäksi he ovat määritelleet ”itsenäisen kirjallisuuskatsauksen”. Itsenäisessä kirjallisuuskatsauksessa on kyse artikkelista, jossa arvioidaan aiemmin julkaistuja tutkimuksia ilman, että artikkeleita varten olisi erikseen kerätty tutkimusaineistoa. Kun itsenäinen kirjallisuuskatsaus toteutetaan systemaattisia ja täsmällisiä standardeja noudattaen, sitä kutsutaan systemaattiseksi kirjallisuuskatsaukseksi (*systematic literature review*, SLR). (Okoli & Schabram, 2010.) Oman graduni sijoittaisin Okolin ja Schabramin määrittelyissä sekä toiseen että kolmanteen ryhmään, sillä gradussa on teoreettista taustaa avaava kirjallisuuskatsaus, mutta myös itsenäinen kirjallisuuskatsaus, jonka tarkoitus on arvioida aiheesta toteutettuja tutkimuksia.

Finkin (2019) mukaan kirjallisuuskatsaus on systemaattinen, täsmällinen ja toistettava menetelmä aiemmin toteutettujen tutkimusten tunnistamiseksi, arvioimiseksi ja kokoamiseksi. Fink (2019) jakaa kirjallisuuskatsauksen toteutuksen seitsemään vaiheeseen:

1. Tutkimuskysymyksen määrittely
2. Tietokannan valinta
3. Hakutermien määrittely
4. Käytännöllisten seulontakriteereiden (esim. julkaisu vuosi, julkaisukieli, artikkelin tyyppi, rahoituslähde) määrittely
5. Metodologisten seulontakriteereiden (tieteellinen laatu) määrittely
6. Arvioinnin toteuttaminen

7. Tulosten yhdistely kuvailevasti

Kirjallisuuskatsaus menetelmänä voidaan jakaa Snyderin (2019) tapaan esimerkiksi systemaattiseen, puolisystemaattiseen ja integroivaan, jotka eroavat toisistaan muun muassa tarkkuuden, laajuuden ja hakustrategian ja analysoinnin perusteella. Systemaattinen kirjallisuuskatsaus on kaikkein täsmällisin ja perusteellisin tapa artikkeleiden kokoamiseen, sillä siinä on tarkoitus pyrkiä löytämään tutkimuskysymyksen kannalta oleellinen tieto. Tämä edellyttää tarkkaan määriteltyä tutkimuskysymystä sekä standardin mukaisten hakustrategioiden ja seurantakriteereiden noudattamista. Puolisystemaattisessa tai puolijärjestelmällisessä kirjallisuuskatsauksessa sen sijaan tutkija yleensä määrittelee itse tutkimuskysymyksen ja prosessin läpinäkyvyyden kannalta sopivat hakustandardit. (Snyder, 2019.)

Valitsin tämän tutkimuksen menetelmäksi systemaattisen kirjallisuuskatsauksen varmistaakseni tutkimuksen luotettavuuden ja läpinäkyvyyden. Systemaattisen kirjallisuuskatsauksen etuna on, että suunnitteluprosessi ja kirjallisuuden tunnistaminen tapahtuvat selkeästi määriteltyjen periaatteiden mukaisesti. Seuraavissa kappaleissa kuvaan, miten olen käytännössä toteuttanut katsauksen suunnittelun sekä kirjallisuuden tunnistamisen, syntetisoinnin ja analysoinnin Finkin määrittelemien seitsemän vaiheen mukaisesti.

3.2 Tietokannan valinta

Kirjallisuushaun tietokantana toimi scholars.google.fi. Finkin (2010) mukaan kattavimmat kokoelmat artikkeleita löytyvät verkon tietokannoista. Verkkoyhteyden päässä ovat koko maailman tieteelliset kirjalliset aineistot ja muut tietolähteet. (Fink, 2010.) Itse koen että ajankohtaisen tiedon etsinnässä verkkotietokannat ovat tänä päivänä ainoa varteenotettava vaihtoehto artikkeleiden hakuun. Mikäli kirjallisuuskatsauksessa haettaisiin historiatietoa, silloin muidenkin tietokantojen käyttö voisi olla perusteltua.

Googlen hakukoneeseen päädyin omassa tutkimuksessani, koska älykkyyttä, tekoälyä ja ongelmanratkaisua tarkastellaan monen ei-tieteenalan näkökulmasta. Uskon, että scholarista löytyy kattavasti tutkimusongelmaani käsitteleviä artikkeleita. Vaikka oma lähestymiseni on kognitiotieteellinen, ja olisin voinut valita hakuprosessiin myös jonkin kognitiotieteellisen tietokannan, koen että riskinä olisi voinut olla suppeampi näkökulma tutkimuskysymykseeni. [Scholars.google](https://scholars.google.com/) -haun kautta oli mahdollista saada mukaan myös monitieteisiä tutkimuksia sekä artikkeleita, joita ei vielä ole virallisesti julkaistu.

3.3 Hakutermien määrittely

Hakutermien huolellinen määrittely on kirjallisuuskatsauksen lopputuloksen kannalta erittäin keskeistä, sillä esimerkiksi haku "*artificial intelligence and*

complex problem solving” tuotti google scholarsissa 18.4.2023 noin 3 230 000 hakutulosta. Löytääkseen tutkimuksen kannalta oleellista kirjallisuutta, eikä vain rajatonta määrää aihetta sivuavia artikkeleita, tutkimustarpeet tulee määritellä tarkkaan (Fink, 2010). Rajatakseni hakua ja löytääkseni kirjallisuuskatsaukseni kohtuullisen määrän artikkeleita toteutin kirjallisuushaun Taulukko 1. kuvatuilla hakutermeillä. Toteutin haun ainoastaan englanninkielisillä hakutermeillä, koska alan keskeiset julkaisut ovat englanniksi. Halusin löytää erityisesti artikkeleita, jotka keskittyvät käsittelemään tekoälyn ongelmanratkaisumahdollisuuksia monimutkaisessa ympäristössä. Tästä syystä keskeisenä hakuterminä oli *ai complex problem solving*. Pohdin ja kokeilin hakuja myös *ill-structured problem solving* -termeillä, jota ovat omassa haussaan käyttäneet mm. Johnson ym. (2022). Tulkitsin koehakujen perusteella, että termi *complex problem solving* olisi yleisemmin käytössä, joten siksi hyödynsin sitä omassa haussani. Kognitiotieteellisen tutkimusotteeni näkökulmasta halusin erityisesti löytää artikkeleita, joissa olisi mukana kognitiotieteellinen näkökulma, joten siksi hakusanoissa on ensimmäisenä mukana *cognitive science*.

Kokeilin erilaisia hakuja (liite 1⁹), joista useilla hakutuloksia tuli kohtuuttoman paljon. Koehakujen perusteella huomasin, että tuhansien hakutulosten määrää sai rajattua erityisesti hyödyntämällä tarkennetussa haussa kohtaa *sisältäen tarkan ilmauksen*. Tarkensin hakua kokeilemalla erilaisia hakukriteereitä ja siten pääsin pikkuhiljaa lähemmäs lopulliseksi muotoutunutta hakua. Lopullisten hakukriteereiden osalta (Taulukko 1) tein ensin koehaun, jonka osalta silmäilin vastaako hakutulos tarpeitani, ennen kuin toteutin varsinaisen haun. Snyderin (2019) mukaan sopivan otoskoon tavoitteluun liittyy yleensä riski siitä, että hakua rajataan liian paljon esimerkiksi ottamalla tarkasteluun vain kapea aikaväli tai rajamalla pois muiden alojen tutkimuksia. Oman hakuni osalta olen pyrkinyt siihen, että rajaus ei olisi liian karsiva, mutta hakutulosten määrä pysyisi kuitenkin kohtuullisena.

Taulukko 1. Hakutermit

Etsi artikkeleita kaikilla sanoilla	cognitive science solutions for ai complex problem solving
sisältäen tarkan ilmauksen	artificial intelligence problem solving
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-

⁹ Liite 1 ei ole kattava listaus kaikista kokeilemistani hauista, vaan sisältää muutamia esimerkkejä kokeilemistani hauista sekä niiden tuottaman hakutulosten määrän.

Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä	2017–2023
---	-----------

3.4 Seulontakriteerit

Kävin läpi kaikki tarkennetun haun kautta saamani artikkelit ja seuloin ne tiivistelmien ja julkaisutietojen pohjalta siten, että jätin synteestistä pois artikkelit, jotka eivät täyttäneet seuraavia kriteereitä:

- Tieteellinen julkaisu
- Julkaisukieli englanti
- Julkaisuvuosi 2017–2023
- Artikkelit saatavilla vapaasti tai yliopiston kautta koko tekstinä
- Artikkelissa käsitellään tekoälyn ongelmanratkaisua monimutkaisten ongelmien näkökulmasta
- Artikkelissa on tunnistettu tekoälyn ongelmanratkaisuun liittyviä haasteita
- Artikkelissa esitetään jokin ratkaisuehdotus tekoälyn ongelmanratkaisukyvyyn kehittämiseksi

Sisällytin hakutulokseen ainoastaan englanninkielisiä artikkeleita, sillä kielitaitoni ei riitä muiden kuin englannin ja suomenkielisten artikkeleiden syvälliseen analysoimiseen. Yhtään suomenkielistä artikkelia ei toteuttamalla haulla tullut mukaan hakutulokseen.

Lähtökohtaisesti kirjallisuuskatsauksissa suositaan yleensä vertaisarvioituja ja julkaistua artikkeleita, mutta tekoälyn osalta laadukkaita ja ajankohtaisia artikkeleita julkaistaan myös arxiv-palvelussa käsikirjoitusmuodossa ja niitä ei välttämättä ole edes tarkoitus julkaista sen jälkeen muualla. Koska arxiv-palvelussa voi olla tutkimuskysymyksen kannalta oleellisia artikkeleita, en ole asettanut seulontakriteeriksi, että artikkeleiden tulisi olla julkaistuja ja vertaisarvioituja. Fink (2010) nostaa julkaisemattomien artikkeleiden osalta esille myös tärkeän näkökulman siitä, että osaa artikkeleista ei julkaista, koska niiden johtopäätökset voivat olla tavanomaisia tai jopa negatiivisia. Julkaisuun päätyy helpommin sellaisia artikkeleista, joissa on saatu positiivisia tai muuten kiinnostavia tuloksia. Ikävimmillään tämä saattaa johtaa virheellisiin tulkintoihin tietystä aihepiiristä. (Fink, 2010.)

Artikkeleiden julkaisuajankohdaksi olen määritellyt vuodet 2017–2023, koska haluan keskittyä erityisesti viimeaikaiseen tutkimustietoon tutkimusongelmani osalta. Vuosi 2017 on valikoitunut aloitusvuodeksi, sillä vuonna 2016 AlphaGo päihitti maailman parhaan Go-pelaajan, mikä oli tekoälyn kehityksessä merkittävä saavutus kahdesta syystä: 1) Go-peli on huomattavasti shakkia monimutkaisempi peli, joka edellyttää strategista suunnittelua ja 2) AlphaGo:n ”älykyys” perustui aiempaa edistyneempiin algoritmeihin ja tekoälyn itseoppiseen (Shushan & Xiaodi, 2017).

Sisällytän mukaan artikkelit, joissa käsitellään selkeästi tekoälyn ongelmanratkaisukykyä ja pohditaan tai esitetään mahdollisia ratkaisuja tekoälyn kehittämiseksi.

3.5 Hakutulosten seulonta

Toteutin kirjallisuuskatsauksen haun 11.5.2023. Kappaleessa 3.3 kuvattujen hakutermien mukainen haku tuotti 48 hakutulosta (Liite 3). Tämän jälkeen kirjasin kaikista hakutuloksista Exceliin Taulukko 2. mukaiset perustiedot.

Taulukko 2. Hakutulosten perustiedot

Linkki	Tyyppi	Otsikko	Tekijät	Julkaisu- vuosi	Julkai- sija	Saata- vuus	Kieli	Ongel- manrat- kaisu	Haas- teet	Ratkai- suehdo- tukset

Perustietojen kirjaamisen jälkeen identifioin ja rajasin hakutuloksen PRISMA-mallia (kuviokuva 6) noudattaen seuraavien kriteereiden perusteella:

- Tieteellinen julkaisu
- Julkaisukieli englanti
- Julkaisuvuosi 2017–2023
- Artikkelit saatavilla vapaasti tai yliopiston kautta koko tekstinä

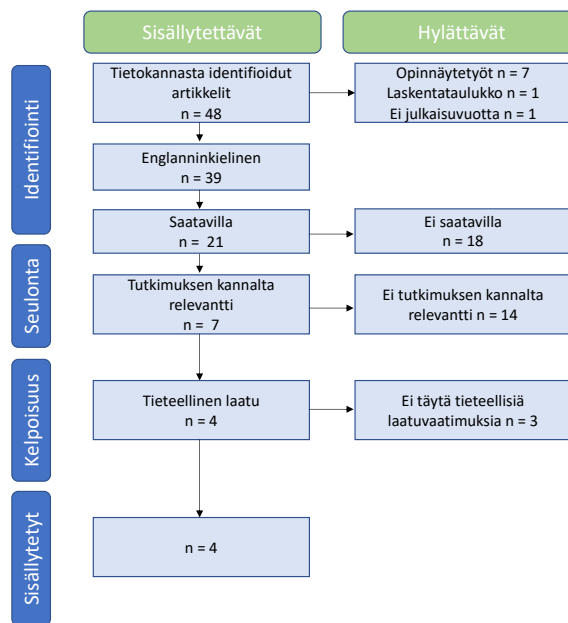
Hakutulosten käsittelyn toiseen vaiheeseen eli seulontaan selvisi 21 artikkelia. Finkin (2010) mielestä tehokas kirjallisuushaku toteutetaan aina hyödyntämällä kaksiosaista seulontaa; ensimmäinen on käytännöllinen tarkastelu ja jälkimmäinen metodologisen laadun arviointi. Ensimmäisessä seulontavaiheessa poimitaan artikkelit, jotka ovat asiaankuuluvia eli ne ovat aiheen otsikon mukaisia, ovat tarkastelukiellellä kirjoitettuja ja ne on julkaistu asianmukaisesti. Toisessa seulontavaiheessa pyritään löytämään sellaiset tutkimukset, joissa on noudatettu tieteellisten tutkimusmenetelmien vaatimuksia ja tuotettu luotettavaa näyttöä. Molempien seulontavaiheiden avulla varmistetaan asiaankuuluva ja virheetön arviointi. (Fink, 2010.)

Seulonnan toteutin arvioimalla otsikon, tiivistelmän ja johtopäätösten perusteella, mitkä artikkeleista käsittelevät tutkimuskysymykseni mukaisia teemoja:

- Artikkelissa käsitellään tekoälyn ongelmanratkaisua monimutkaisten ongelmien näkökulmasta
- Artikkelissa on tunnistettu tekoälyn ongelmanratkaisuun liittyviä haasteita
- Artikkelissa esitetään jokin ratkaisuehdotus tekoälyn ongelmanratkaisukyvyn kehittämiseksi

Tutkimuksen kannalta relevanteiksi tunnistin sellaiset artikkelit, jotka täyttivät kaikki kolme edellä mainittua kriteeriä. Rajasin pois muun muassa sellaiset artikkelit, joissa ei esitetty ratkaisuehdotuksia tekoälyn ongelmanratkaisukyvyyn kehittämiseksi.

Tutkimusten tieteellisen laadun kartoittamiseksi pyrin tiivistelmän ja johtopäätösten pohjalta arvioimaan tutkimusten johdonmukaisuutta ja laatua. Lisäksi selvitin julkaisujen luokittelut julkaisufoorumissa. Tarkempi kuvaus tieteellisen laadun arvioinnista Liitteessä 2. Tieteellisen laadun arvioinnin jälkeen hakutulokseen sisällytettäväksi jäi neljä artikkelia.

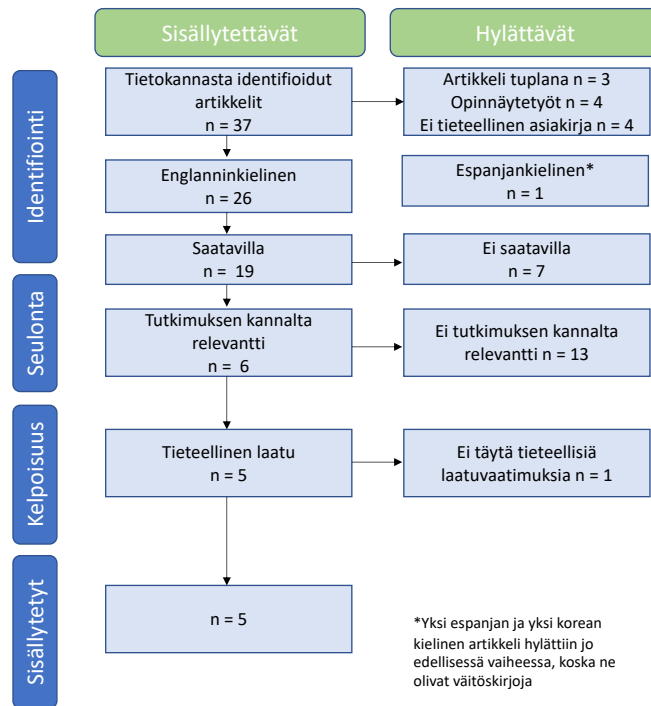


Kuvio 6. PRISMA-mallia mukaileva kuvaus hakutulosten seulonnasta

Koska hakuni tuotti ainoastaan neljä seulonta- ja kelpoisuuskriteerit täyttävää hakutulosta, päätin tehdä vielä täydentävän haun. Puuttuvien tietojen tunnistaminen ja tarvittaessa tietojen täydentäminen kuuluvat Hirsjärvi, Remes ja Sajavaara (1997) mukaan aineiston käsittelyn esitöihin. Tarkasteltuani hakutuloksiani päätin tehdä toisen haun ilman *cognitive science* -termiä. Kyseisen hakusanan jättäminen pois oli mielestäni perusteltua siitä syystä, että se ei ollut tuottanut hakutuloksiini toivottua vaikutusta kognitiotieteellisen tutkimusnäkökulman korostumisesta, vaan hakutuloksissa oli laajasti eri tieteenalojen artikkeleita. Koska hakusanalla ei ollut selkeää vaikutusta, oli se perusteltua jättää pois toisesta hausta.

Toisen haun toteutin 19.5.2023 muuten samoilla hakukriteereillä, kuin ensimmäisellä kerralla, mutta ilman *cognitive science* -termiä. Tällä tavalla sain 86 hakutulosta. Hakutulokseen sisältyivät kaikki ensimmäisellä haulla tunnistetut hakutulokset, kolme artikkelia oli tuplana ja yksi hakutuloksista oli pelkkä sitaatti. Duplikaattien, sitaatin ja aiemmin seulottujen artikkeleiden poiston jälkeen jäljelle jäi 34 uutta artikkelia (Liite 4), jotka kävin läpi vastaavalla tavalla kuin ensimmäisellä kerralla (kuvio 7). Toisen kierroksen haun osalta jäljelle jäi

viisi hakutulosta. Yhteensä kirjallisuushakuni tuotti siis kahdeksan tutkimukseeni sisällytettävää artikkelia.



Kuvio 7. PRISMA-mallia mukaileva kuvaus toisen kierroksen hakutulosten seulonnasta

3.6 Aineiston analyysi

Snyder (2019) mukaan lopullisen otannan valinnan jälkeen tulee hyödyntää standardoitua tapaa oleellisen informaation tunnistamiseksi jokaisesta artikkelista. Okoli ja Schabram (2010) mukaan jokaista artikkelia tulisi tarkastella neljästä näkökulmasta: Mitä väitteitä esitetään, mitä todisteita väitteille esitetään, ovatko todisteet perusteltuja ja miten ne on taustoitettu. Omassa tutkimuksessani toteutin tämän siten, että luin läpi kaikki seulontakriteerit täyttäneet kahdeksan artikkelia kirjatun niistä Excel-taulukkoon (ks. Taulukko 3) seuraavat tiedot:

- Tutkimuksen tavoite
- Tutkimusmenetelmä
- Tekoälyn ongelmanratkaisukyvyyn osalta tunnistetut keskeiset ongelmat
- Keskeiset tutkimustulokset
- Ratkaisuehdotukset, jotta tekoäly kykenisi ongelmanratkaisuun monimutkaisissa ja jäsentymättömissä tilanteissa

- Huomioita¹⁰

Excel-taulukkoon kirjaamani tiedot toimivat siis tutkimukseni analyysiyksiköinä eli ajatuskokonaisuuksina, jotka vastaavat esitettyihin tutkimuskysymyksiin (Elo ym., 2022). Aineiston analyysin yleisenä tavoitteena on arvioida aineistoa kriittisesti ja tutkia artikkeleiden keskeisiä ajatuksia ja niiden välisiä keskinäisiä riippuvuuksia (Snyder, 2019). Tässä tutkimuksessa olen tarkastellut ja analysoinut aineistoa erityisesti tutkimuskysymysteni näkökulmasta.

3.7 Validiteetti

Tutkimuksen validiteetin arviointi on Hayashi ym. (2019) mukaan tärkeä tapa todentaa tutkimuksen laadukkuus, mutta kvalitatiivisissa tutkimuksissa sen toteuttamiseen ei ole kvantitatiivisen tutkimuksen tapaan yleisesti määriteltyjä kriteereitä. Snyder (2019) suosittelee, että kirjallisuuskatsauksen laatua tarkasteltaisiin suhteessa jokaisen vaiheen toteutukseen: Suunnittelu, toteutus, tiedon käsittely ja analysointi sekä arvioinnin jäsentely ja kirjoittaminen. Hayashi ym. (2019) ovat tunnistaneeet kirjallisuudessa useita erilaisia tapoja kvalitatiivisen tutkimuksen laadun arviointiin ja validiteetin luokitteluun, mutta heidän mukaansa validiteetin tulisi ensisijaisesti rakentua osana koko tutkimusprosessia ja olla osa jokaista tutkimuksen osa-aluetta. Hayashi ym. (2019) mukaan prosessuaalisen validiteetin osa-alueita ovat A) Tutkimuksen laajuuden, menetelmien ja rajojen määrittely B) Tiedon hankinnan organisointi C) Tiedon kokoaminen ja analysointi D) Tiedon analysointi osana laajempaa ilmiötä E) Tulosten tulkinta. Mielestäni Snyder (2019) ja Hayashi ym. (2019) esittämät jäsentelyt muistuttavat vahvasti toisiaan. Koska kvantitatiivisen tutkimuksen validiteetin arvioinnin perusteet soveltuvat epistemologisten ja ontologisten erojen vuoksi huonosti laadulliseen tutkimukseen, olen seuraavissa kappaleissa arvioinut oman tutkimukseni validiteettia hyödyntäen Snyder (2019) ja Hayashi ym. (2019) mukaista rakennetta. Suhteutettuna kvantitatiivisen tutkimuksen validiteetin arviointiin suunnittelu ja toteutus vastaavat mielestäni sisäistä validiteettia, tiedon käsittely ja analysointi rakennevaliditeettia ja tiedon kokoaminen ulkoista validiteettia. Ekologisen validiteetin näkökulma linkittyy useampaan osaan, mutta tulee parhaiten esille tutkimuksen toteutuksen osuudessa.

3.7.1 Suunnittelun validiteetti

Snyder (2019) mukaan suunnittelun validiteettia tarkastellessa tulee pohtia muun muassa tarvitaanko tätä kirjallisuuskatsausta, ovatko tutkimus-

¹⁰ Nämä ovat artikkelista heränneistä omia ajatuksiani, jotka koin tarpeellisiksi kirjata muistiin analysointia tai johtopäätösten tekemistä ajatellen. Elo ym. (2022) perusteella huomioita voitaisiin nimittää myös koodauksen osaksi eli aineistoa jäsentevien merkintöjen tekemiseksi.

kysymykset selkeästi perusteltu, miten valittu menetelmä vaikuttaa tutkimuksen luotettavuuteen ja onko valittu menetelmä perusteltu ja selkeästi kuvattu.

Tekoölyn ongelmanratkaisun tutkimukselle on mielestäni selkeät perusteet ja tarve. Tekoölyn ongelmanratkaisukyvyyn ongelmien tunnistaminen ja ratkaisujen arviointi auttavat kehittämään järjestelmiä, jotka tekevät luotettavampia, toimivampia ja läpinäkyvämpiä ratkaisuja yhä monimutkaisemmissa tilanteissa.

Tutkimuskysymyksenäni ovat laajat minkä tiedostin jo tutkimuksen alkuvaiheessa. Halusin kuitenkin luoda yleiskatsauksen tekoölyn ongelmanratkaisukykyyn suhteessa ihmisen kykyihin ja luoda siten synteisiä tulevia tutkimuksia varten. Tarkemmin rajatuilla kysymyksillä olisin voinut perehtyä syvällisemmin esimerkiksi tiettyihin tekoölyn kehittämisen ratkaisuihin, mutta tällöin kokonaisvaltaisen kehittämisen näkökulma olisi voinut jäädä puuttumaan.

Tutkimusmenetelmäksi oli luontevaa valita systemaattinen kirjallisuuskatsaus, koska minulla ei ollut etukäteen tiedossa toimivaa ratkaisua, jota olisin voinut lähteä arvioimaan ja testaamaan. Tiesin kirjallisuuden pohjalta myös, ettei tekoöly vielä kykene ongelmanratkaisuun monimutkaisissa ja jäsentymättömissä tilanteissa, joten empiirinen testaaminen ei olisi onnistunut: Sellaista ratkaisua, jota ei ole olemassa, on vaikea testata. Tutkimuskenttää luotaavia menetelmiä olisivat toki voineet olla myös asiantuntijoiden haastattelut sekä kyselytutkimus tekoölytutkimusta tehneille henkilöille tms. Toisaalta tekoölytutkimuksen valtava määrä mielestäni edellyttää myös sitä, että julkaistuista artikkeleista laaditaan synteesejä ja arvioidaan kriittisesti niissä esitetyt ratkaisut. Kirjallisuuskatsauksen menetelmänä toki liittyy aina riski siitä, että haulla ei onnistuta löytämään tutkimuskysymysten kannalta keskeisimpiä tutkimuksia. Tätä ongelmaa olen pyrkinyt minimoimaan muun muassa systemaattisen haun sekä hakusanojen testaamisen myötä.

Mielestäni systemaattinen kirjallisuuskatsaus oli perusteltu tutkimusmenetelmä ja olen pyrkinyt kuvaamaan selkeästi ja läpinäkyvästi, miten sitä on hyödynnetty. Tutkimusongelma, teoria, tutkimuksen toteutus ja tutkimustulokset muodostavat johdonmukaisen ja yhtenäisen kokonaisuuden.

3.7.2 Toteutuksen validiteetti

Toteutuksen validiteetin tarkastelussa Snyder (2019) ohjeistaa pohtimaan muun muassa hakusanojen vaikutusta, hakuprosessin toteutusta, otoksen edustavuutta sekä toteutuksen laatua. Olen kappaleessa 3 pyrkinyt kuvaamaan tutkimuksen toteutuksen mahdollisimman läpinäkyvästi ja tarkasti huolehtien siten siitä, että vastaava hakuprosessi olisi kenen tahansa toteutettavissa. Hakuprosessin olen kuvannut PRISMA-vuokaavion pohjalta, joka on yleisesti käytössä systemaattisten kirjallisuuskatsausten dokumentoinnissa ja osaltaan vahvistaa tutkimuksen toistettavuutta ja siten edistää validiteettia.

Valitsemillani hakusanoilla on ollut merkittävä vaikutus siihen, minkälaisia hakutuloksia olen kirjallisuuskatsaukseni saanut. Vaikka tein useita testihakuja erilaisilla hakusanayhdistelmillä, optimaalisen hakutuloksen saavuttamiseksi sopivia hakusanoja olisi voinut selvittää ja kokeilla vielä kattavammin. Hirsjärvi ym. (1997) tosin huomauttavat, että tutkijalle voi vasta analyysivaiheessakin selvitä, miten ongelmat olisi pitänyt asettaa. Myös omalla kohdallani ymmärsin

vasta tutkimuksen viimeistelyvaiheessa, että käyttämäni hakukriteeriin *sisältää tarkan ilmauksen artificial intelligence problem solving* liittyy validiteettiongelma. En ollut hakutermejä muotoillessani tullut ajatelleeksi, etteivät kyseiset sanat varsinaisesti esiinny luonnostaan tekstissä käyttämässäni muodossa. Käytännössä olen siis saattanut kyseisellä hakukriteerillä rajata ulos tutkimuskysymysteni kannalta oleellisia hakutuloksia. Haun rajaaminen sopivan kokoiseksi tuottaa helposti tilanteen, jossa hakusanojen määrittelyssä joutuu tekemään tiettyjä kompromisseja. Valitsemillani hakusanoilla löysin hallittavan määrän artikkeleita, joissa käsitellään molempia tutkimuskysymyksiäni, mutta ne eivät välttämättä ole kattava otos aineistosta. Toisaalta ilman tarkan ilmauksen kriteeriä haku olisi tuottanut noin 19 000 hakutulosta. Tarkan ilmauksen sisällyttäminen oli siis perusteltua, mutta jälkikäteen arvioituna sen muotoilua olisi pitänyt tarkastella perusteellisemmin.

Lähestyn tekoälyn ongelmanratkaisua kognitiotieteellisestä näkökulmasta ja toivoin löytäväni kirjallisuuskatsaukseen aineistoa, jossa ongelmanratkaisua tarkasteltaisiin tästä näkökulmasta. Hakukriteerini täyttäneiden artikkeleiden osalta ainoastaan yhdessä oli mielestäni kognitiotieteeseen linkittyvä näkökulma ja sekään ei tulkintani mukaan ollut kognitiotieteellinen tutkimus. Tämä riski oli olemassa, kun toteutin haun google.scholarsin kautta. Mikäli olisin toteuttanut haun kognitiotieteellisiin julkaisuihin keskittyneen hakukoneen kautta, olisin oletettavasti saanut kattavammin kognitiotieteellisen näkökulman mukaisia artikkeleita. Tällöin hakutulos olisi kuitenkin saattanut jäädä suppeammaksi ja näkökulma tekoälyn kehittämisen tarpeisiin olla kapeampi. Tällöin myös validiteetti olisi heikentynyt esimerkiksi triangulaation näkökulmasta (ks. 3.7.3). Yritin löytää google.scholarsista kognitiotieteeseen liittyviä artikkeleita hyödyntämällä haussa cognitive science käsitteitä. Huomasin kuitenkin hakutuloksia käsitellessäni, että tällä ei ollut vaikutusta, joten toteutin uuden haun ilman kyseisiä hakutermejä ja sain näin hieman laajennettua hakutulostani. Muutos ei karsinut pois yhtään ensimmäisellä hakukerralla tunnistettua artikkeleita, mutta toi tarkasteluun 38 artikkelia lisää. Mielestäni hakusanojen muokkaus ensimmäisen haun jälkeen oli siis perusteltua.

Otoskoon sopivuutta voidaan tarkastella muun muassa saturaation näkökulmasta. Hayashi ym. (2019) mielestä saturaatio saavutetaan, kun aineistosta ei nouse enää uutta hyödyllistä informaatiota. Oman aineistoni osalta en voi todeta saturaation täytyneen sillä artikkeleissa ei käsitelty kattavasti kaikkia tekoälyn ongelmanratkaisukykyyn liittyviä ongelmia ja toisaalta esitetyt ratkaisuehdotukset olivat kaikissa erilaisia. Otoksen ei siis voida katsoa olevan täysin kattava katsaus tekoälyn ongelmanratkaisukykyyn. Tämän tutkimuksen laajuus ja käytettävissä olevat tutkimusresurssit huomioiden suurempi otos olisi saattanut kuitenkin tuoda ongelmia aineiston hallittavuuden näkökulmasta. Mielestäni yksin toimivan tutkijan ei ole mahdollista kohtuullisessa ajassa käydä läpi valtavaa aineistoa, koska aineiston identifiointi, seulonta ja analyysi vaativat tarkkuutta ja huolellisuutta. Tarkan otoskoon määrittelyn sijaan kirjallisuuskatsauksissa pyritään kiinnittämään huomiota hakuprosessin ja seulonnan johdonmukaiseen rakenteeseen, ymmärrettävyyteen ja läpinäkyvyyteen (Hiebl, 2023).

Yhtenä keskeisenä näkökulmana kirjallisuuskatsauksen validiteetin arvioinnissa ovat periaatteet, joilla katsaukseen sisällytetään tai suljetaan

ulkopuolelle artikkeleita. Snyder (2019) korostaa, että tutkijan tulisi kiinnittää erityistä huomiota siihen, mitä tutkimuksia katsaukseen sisällytetään ja millä perusteilla, koska nämä valinnat vaikuttavat olennaisesti siihen, mitä johtopäätöksiä saavutetaan. Oleellisia tutkimusaloja, artikkeleita tai tutkimusvuosia poissulkemalla on merkittäviä vaikutuksia tutkimuksen tuloksiin ja johtopäätöksiin. Myös Fink (2010) muistuttaa, että kaikki tutkimukset eivät ole yhtä laadukkaita ja arvioijan tulee pystyä tunnistamaan hyvä- ja huonolaatuiset tutkimukset. Korkealaatuisen tutkimuksen tavoitteena on tuottaa virheetöntä tietoa ja jos kirjallisuuskatsauksen perustana on ei niin korkealaatuista tutkimusta, myös tulokset ovat vähemmän paikkansa pitäviä. (Fink, 2010.) Omasta näkökulmastani olen toteuttanut aineiston seulonnan huolellisesti ja kriteerit artikkeleiden sisällyttämiseen ja pois sulkemiseen on kuvattu täsmällisesti. Seulontaprosessi eteni kappaleessa 3.5 kuvatun mukaisesti, eikä siinä ilmennyt ristiriitoja tai epä johdonmukaisuuksia. Ainoastaan yhden kirjan osalta jouduin tekemään syvällisemmän pohdinnan pitäisikö se sisällyttää aineistoon. Ongelmanratkaisua käsiteltiin kirjassa useammassa kohdassa, mutta en pystynyt identifioimaan, minkälaisia ratkaisuehdotuksia tekoälyn ongelmanratkaisukyvyyn kehittämiseksi kirjassa esitetään tai esitetäänkö varsinaisesti mitään uusia ratkaisuja, joten päädyin lopulta rajaamaan kirjan pois hakutuloksista. Mikäli kirjallisuuskatsaukseen olisi halunnut vielä vahvemmin kognitiotieteellisen näkökulman, yhtenä seulontakriteerinä olisi voinut olla esimerkiksi jokin tämän kaltainen kriteeri: ratkaisuehdotus on johdettu ihmisen älykkyyden pohjalta. Tällaisella kriteerillä lopullinen tutkimusaineisto tosin olisi kutistunut merkittävästi ja saattanut jäädä tarkastelunäkökulmilta liian kapeaksi.

Kirjallisuuskatsauksessa hakutulokseen vaikuttaa myös aineiston saatavuus, mikä Norris (1997) mukaan on yksi validiteetin näkökulma. Kaikki tutkimuskysymyksen kannalta oleellinen aineisto ei välttämättä ole ollut saatavilla hakukoneiden kautta tai se ei ole ollut saatavilla vapaasti tai yliopiston kautta koko tekstinä, mikä oli yksi seulontakriteereistäni. Osa potentiaalisesta aineistosta on siis karsiutunut pois siitä syystä, että ne eivät ole olleet saatavilla.

3.7.3 Tiedon käsittelyn ja analysoinnin validiteetti

Hayashi ym. (2019) mukaan laadullisen tutkimuksen parhaiten tunnettu validiteetin arvioinnin kriteeri on triangulaatio (*triangulation*). Triangulaatio muodostuu aineistosta kerätyn tiedon ja tutkimuskysymysten välisen suhteen pohjalta ja sen tavoitteena on vahvistaa tulosten reliabiliteettia. (Hayashi ym. 2019.) Käytännössä triangulaatio tarkoittaa useiden eri tietoaineistojen, menetelmien, teorioiden tai tutkijoiden hyödyntämistä tutkimuskysymykseen vastaamiseksi (Bhandari, 2022). Tässä tutkimuksessa on hyödynnetty vain yhtä menetelmää ja tutkijoita on ollut yksi, joten niiltä osin triangulaatio on heikko. Teorioiden ja tietoaineistojen näkökulmasta tämän tutkimuksen triangulaatiota voidaan kuitenkin tarkastella laajemmin. Bhandari (2022) esittää, että aineiston kohdalla triangulaatiota voidaan toteuttaa hankkimalla aineistoa eri aikoina, eri henkilöiltä ja eri paikoista. Omassa kirjallisuuskatsauksessani tämä toteutuu mielestäni siten, että katsaukseen seulottu aineisto on julkaistu eri kanavissa, eri aikoina ja eri henkilöiden toimesta. Tämä on osittain ollut sattumanvaraista, sillä hakua ei ole

suoraan ohjattu tuottamaan määriteltyä vaihtelua. Toisaalta valitsemalla hakukoneeksi google.scholarsin olen edistänyt julkaisukanavien monipuolisuuden toteutumista ja määrittämällä aikajänteeksi 2017–2023 olen saanut aineistooni eri aikoina julkaistuja artikkeleita.

Teorian triangulaatio tarkoittaa Bhandari (2022) mukaan sitä, että tutkimuskysymystä lähestytään useiden teoreettisen viitekehysten näkökulmasta. Oman tutkimukseni kohdalla teoreettinen triangulaatio on mielestäni kiinnostava kysymys. Näkökulmani on kognitiotieteellinen, mutta tekoälytutkimus itsessään on jo niin monitieteistä (mm. psykologia, matematiikka, tietojenkäsittelytiede, neurotiede, kognitiotiede), että sitä olisi uskoakseni lähes mahdoton tarkastella pelkästään yhden tieteenalan kontekstista. Toisaalta tutkimuskysymykseni on asetettu kognitiotieteellisestä näkökulmasta ja tarkastelen aineistoa myös siltä kannalta. Aineisto sinänsä ei ole pelkästään kognitiotieteellistä, sillä hyödyntämällä google.scholarsia hakukoneena, olen saanut eri lähteissä julkaistuja eri tieteenaloilla julkaistuja artikkeleita. Nähdäkseni teorian triangulaatio toteutuu siis varsin hyvin.

Kuten hakutulosten seulonnan, myös tietojen kokoamisen olen kuvannut tarkasti ja johdonmukaisesti. Taulukkoon 3 olen koonnut keskeiset tiedot tutkimusaineistosta. Kyseinen taulukko kuvastaa myös minkälaisilla periaatteilla tiedon kokoaminen on tapahtunut. Tiedon kokoamisessa olen pyrkinyt johdonmukaiseen ja yhtenäiseen käsittelyyn, mutta useita artikkeleita lukiessa tutkijan väsyminen on mahdollista. Väsymisen vaikutuksia olen pyrkinyt vähentämään useiden käsittelykertojen kautta. Ensin olen lukenut kaikista artikkeleista tiivistelmän ja johtopäätökset, tämän jälkeen olen lukenut kaikki artikkelit kertaalleen läpi ja kolmannella lukukerralla olen pyrkinyt analysoimaan ja ymmärtämään sisältöä syvällisesti. En ole käsitellyt artikkeleita aakkosjärjestyksessä, vaan olen edennyt satunnaisessa järjestyksessä eli olen eri kerroilla aloittanut lukemisen eri artikkeleista. Mikäli olisin perehtynyt artikkeleihin aina samassa järjestyksessä, olisi viimeisen artikkelin kohdalla ollut erityinen riski tutkijan väsymyksestä johtuvaan vähäisempään huomioon. Tutkijan väsymyksen vaikutuksia olisi voinut myös vähentää hyödyntämällä useampia tutkijoita tiedon kokoamisessa. Tässä tutkimuksessa useamman tutkijan hyödyntäminen ei kuitenkaan ollut mahdollista.

Aineiston analyysin olen toteuttanut tutkimuskysymysteni näkökulmasta. Tunnistamalla ja tulkitsemalla aineistossa ilmeneviä teemoja olen pyrkinyt syvällisen ymmärryksen muodostamiseen tutkittavasta aiheesta. Tiedon käsittelyn validiteetin osalta tähän liittyy mahdollisen vinouman vaara. Toimiessani yksin tulosten kokoaminen ja analyysi pohjautuvat omiin tulkintoihini ja ajatuksiini, jotka voivat olla vinoutuneita tai puolueellisia. Useamman tutkijan ryhmä olisi voinut tunnistaa monipuolisemmin erilaisia näkökulmia ja toisaalta myös kyseenalaistaa aineistosta tehtyjä tulkintoja. Vähentääkseni mahdollisen vinouman vaikutusta olen yrittänyt kyseenalaistaa omia ajatuksiani ja tunnistaa sellaisia ajatuksia, jotka voisivat ohjata tulkintojani. Olen tunnistanut, että ennako-oletuksenani oli, ettei tekoäly kykene ihmisen kaltaiseen ongelmanratkaisuun. Tämä on varmasti ohjannut ajatteluaani ja tekemiäni valintoja läpi tutkimuksen. Toinen tutkimukseen vaikuttanut tekijä on varmasti ollut kognitiotieteellinen taustani, jonka myötä en esimerkiksi pysty tulkitsemaan kaikilta osin kovin

syvälle teknisiin tai tietojenkäsittelytieteellisiin ratkaisuihin keskittyviä kuvauksia aineistossa. Tämä on saattanut johtaa joiltakin osin vääriin tulkintoihin. Olen kuitenkin pyrkinyt tulkitsemaan ja ymmärtämään aineistoa mahdollisimman syvällisesti sekä välttänyt tulkitsemasta sellaisia asioita, joihin asiantuntemukseni ei riitä.

3.7.4 Tutkimuksen tulosten käsittelyn validiteetti

Snyder (2019) ja Hayashi ym. (2019) olivat nimenneet validiteetin arvioinnin viimeisen vaiheen hieman eri tavoin. Snyder (2019) puhuu arvioinnin jäsentelystä ja kirjoittamisesta, kun taas Hayashi ym. (2019) tulosten tulkinnasta. Käsittääkseni molemmissa on kuitenkin kyse siitä, että validiteettia tarkastellaan sen osalta, miten tulokset on kuvattu suhteessa tutkimuskysymyksiin, minkälainen merkitys tuloksilla on ja ovatko tulokset yleistettävissä.

Tutkimuksen tulosten käsittelyn validiteettia voidaan tarkastella muun muassa tehtyjen valintojen ja rajausten luotettavuuden kautta. Kappaleessa 4 olen käsitellyt tuloksia tutkimuskysymysteni näkökulmasta. Olen jokaisesta artikkelista pyrkinyt tunnistamaan tutkimuskysymysten kannalta keskeiset teemat ja kuvannut ne suhteessa teoreettiseen viitekehykseen. Mielestäni tulosten käsittely on tapahtunut johdonmukaisesti ja avoimesti, taulukoista 4 ja 5 lukijan on helppo havainnoida kokonaisuutta ja tekemiäni tulkintoja. Toisaalta aineiston tulkintaan vaikuttavat aina myös seuraavat tekijät: omat ennakkokäsitykset, tutkijan oma tarkastelunäkökulma, vahvistusharha (*confirmation bias*)¹¹, tutkijan osaaminen ja tutkijan henkilökohtaiset ominaisuudet (mm. McKinnon, 1988; Norris, 1997). Edellä mainittujen tekijöiden negatiivisia vaikutuksia voidaan Norrisin (1997) mukaan vähentää tunnistamalla ja haastamalla omat ennakkosenteet ja oletukset. Myös itsensä havainnointi ja omien tulkintojen kyseenalaistaminen vähentävät tutkijasta johtuvien vinoumien ja virheidensä mahdollisuuksia. (Norris, 1997.) Omalta osaltani olen pyrkinyt suhtautumaan tutkimusaineistosta nouseviin ratkaisuehdotuksiin mahdollisimman avoimesti, mutta samalla tutkimusten laatua kriittisesti arvioiden. Koen onnistuneeni avoimuudessa siinä mielessä, että olen käynyt kriittisesti läpi kaikki hakukriteerit täyttäneet artikkelit, vaikka ne eivät sisältäneetkään enakkotoiveideni mukaista kognitiotieteellistä näkökulmaa. Uskoisin, että samoilla tutkimuskysymyksillä toinenkin tutkija olisi voinut tehdä tutkimusaineistostani vastaavat havainnot. Sen sijaan tuloksia koskevat päätelmät voisivat olla erilaiset riippuen tutkijan kiinnostuksen kohteista ja omista taustoista, esimerkiksi robotiikasta kiinnostunut tutkija olisi saattanut nostaa tulosten keskiöön Mokhtari ym. (2019) EBPD:tä käsittelevän tutkimuksen, kun taas matemaattiselta kannalta asiaa tarkasteleva olisi voinut painottaa tulosten käsittelyssä Jonesin ja Tanimoton (2018) konealbegraa käsittelevää tutkimusta. Vaikka tutkijan tulee aina pyrkiä mahdollisimman kattavaan avoimuuteen ja puolueettomuuteen, mikään ei Norrisin (1997) mukaan takaa

¹¹ Vahvistusharha eli *confirmation bias* tarkoittaa sitä, että tietoa haetaan, tulkitaan ja muistetaan omaa hypoteesia vahvistavasta näkökulmasta (Oswald & Grosjean, 2012).

tutkimuksen virheettömyyttä, tarkoituksenmukaisuutta, totuudellisuutta tai objektiivisuutta.

Oman tutkimukseni tulokset ovat yhdenmukaisia teoreettisen viitekehyksen kanssa, mikä mielestäni osoittaa ainakin jossain määrin niiden olevan luotettavia ja edustavan laajempaa tutkimuksellista näkökulmaa. Tuloksista on myös löydettävissä ratkaisuehdotusten välisiä yhteneväisyyksiä, mikä mielestäni vahvistaa tulosten yleistettävyyttä. Tulokset eivät kuitenkaan ole kattavia, mikä johtuu muun muassa tutkimusongelman laajuudesta sekä kirjallisuuskatsauksen rajallisuudesta. Uskon, että tuloksia voidaan pitää suuntaa antavina.

4 TULOKSET JA SYNTEESI

Snyderin (2019) mielestä kirjallisuuskatsauksissa usein vain luetellaan julkaistujen artikkeleiden määrä, käsitellyt aiheet ja käytetyt menetelmät eli käytännössä ne ovat kuvailevia yhteenvetoja ilman syvällisempää analyysiä. Jotta tämä tutkimus ei jää edellä kuvatun mukaiseksi yleiskatsaukseksi, pyrin analysoimaan aineiston keskeisiä tuloksia molempien tutkimuskysymysteni näkökulmasta. Kuvaan ensin tutkimusaineistoni tiiviissä muodossa, jonka jälkeen käyn läpi minäkälaisia ongelmia tekoälyn ongelmanratkaisun osalta on tunnistettu ja miten ne suhteutuvat muussa kirjallisuudessa tunnistettuihin ongelmiin. Kolmannessa kappaleessa kuvailen, millaisia ratkaisuja tekoälyn ongelmanratkaisun kehittämiseksi esitetään, mitä malleja ja teorioita ehdotetaan ja miten ne vertautuvat alan muihin tutkimuksiin. Pohdin myös, mitkä esitetyistä ratkaisuista vaikuttavat omasta mielestäni lupaavilta. Pyrin referoimaan tutkimuksia objektiivisesti ja arvioimaan niiden ratkaisuehdotuksia kriittisesti.

4.1 Kirjallisuuskatsauksen tiivistelmä ja yleiset huomiot

Tiivistelmä kirjallisuuskatsauksen tuloksena saadusta aineistosta on kuvattu taulukossa 3. Yleisenä huomiona aineistosta voidaan todeta, että vaikka artikkeleiden tarkastelukulmat tekoälyn ongelmanratkaisuun eroavat toisistaan, niissä esitetyt näkökulmat ovat mielestäni yhdenmukaisia ja toisiaan täydentäviä. Artikkeleiden välillä ei ole poikkeavia tai ristiriitaisia näkökulmia. Aineistosta on suhteellisen helposti löydettävissä yhteneväisyyksiä ja muutamia selkeästi toistuvia teemoja, kuten tiedon hankinnan ongelmaan liittyvä mallintamisen vaikeus sekä ajatus ihmisen ja tekoälyn välisestä yhteistyöstä ratkaisuna nykyisiin haasteisiin.

Useammassa artikkelissa mainittiin tavoitteeksi viitekehityksen luominen tutkimusaiheesta. Itselleni heräsi tämän osalta ajatus, kuvaako tarve viitekehityksen luomiseen sitä, että tutkittava aihealue on vielä melko tuore ja siitä saatavilla olevat tiedot ovat vielä kenties rajalliset, jolloin viitekehityksen luomiselle on tarvetta kokonaisuuden jäsentämiseksi.

Muutamassa artikkelissa (Correia ym., 2019; Hwang ym., 2018; X. Li ym., 2020) teorian ja tutkimuksen kuvaus on niin suppea tai monimutkainen, että niistä on vaikea tehdä aukottomia tulkintoja esitettyjen ratkaisujen mahdollisuuksista. Esimerkiksi Hwang ym. (2018) eivät kuvaa selkeästi, mihin tekoälyn ongelmanratkaisun haasteeseen tutkimuksessa esitetään ratkaisua. Myös artikkelin sovellusmahdollisuuksien arviointi on haastavaa, koska asia on kuvattu hyvin suppeasti. Correia ym. (2019) sen sijaan keskittyvät valtaosin viitekehysten esittelyyn, joten joukkoälykkyyden hyödyntämismahdollisuuksien sekä tekoälyn kehittämisen hyötyjen esittely jää suppeaksi tai puuttuu. Li ym. (2020) sen sijaan pohjaavat oman ratkaisuehdotuksena extenics-teoriaan, jota he eivät artikkelissaan juurikaan kuvaa tai määrittele. Olen kuitenkin jättänyt edellä mainitut artikkelit mukaan tarkasteluun, sillä niissä on puutteista huolimatta mielestäni tärkeitä näkökulmia tekoälyn ongelmanratkaisun kannalta. Painotan kuitenkin ratkaisujen esittelyssä niitä artikkeleita, joissa esitetyt teoriat ovat mielestäni huolellisesti ja ymmärrettävästi perusteltu.

Taulukko 3. Tiivistelmä kirjallisuuskatsauksen keskeisistä artikkeleista

Tekijät	Tutkimuksen tavoite	Menetelmä	Tekoälyn ongelmat	Keskeiset tulokset	Ratkaisu-ehdotukset
(Bulgiu ym., 2020)	1) Uusia mittareita määrittämään <i>Blocks World</i> ongelman kokoa ja monimutkaisuutta 2) Ehdottaa algoritmeja mittareiden arviointiin	Mitataan tilojen, siirtymien ja keskiarvoisen haarautumisen määrää <i>Blocks World</i> graafissa hyödyntäen Python listoja ja NumPy ryhmittelyä	Tila-avaruukskaaviot ovat kooltaan valtavia NP (<i>nondeterministic polynomial</i>) ongelmiin ei ole polynomisiin algoritmeihin pohjautuvia ratkaisuja ja <i>Deadlocks</i> vaikuttavat siihen, onko ongelma ratkaistavissa polynomisessa ajassa	Esitetty algoritmi on tehokas hakuavuiden mittaamiseen <i>Blocks World</i> ongelmassa on pariton määrä tiloja ja parillinen määrä siirtymiä	Analyysin tuloksia ja ehdotettuja algoritmeja voitaisiin laajentaa muiden mittareiden sekä tilojen monimutkaisuuden arviointiin (esm. BW4T:ssä)
(Moral ym., 2020)	Uudenlaisten ongelmaratkaisutekniikoiden tuottaminen ja testaaminen Truco-pelin avulla: a) Uusien tilanteiden säilyttäminen b) Uusien tilanteiden säilyttäminen korvaamalla c) Satunnaisten toimien toteuttaminen uusien ongelmaratkaisukokemusten synnyttämiseksi d) Aktiiviseen oppimisstrategiaan pohjautuva tutorointi	Tekniikoita testattiin 300 erilaisessa agenttien välisessä pelitalanteessa	Ongelmien ja ratkaisujen mallintaminen monimutkaiseen peliympäristöön on työläästä CBR liittyy vinouman vaaroja: vastaavaa esimerkkitalannetta ei ole tai tietty toiminta on yliedustettuna opetusmateriaalissa	Agentit suoriutuivat testattavana olleiden uusien oppimistekniikoiden avulla pelistä paremmin, kuin jos käytössä oli ainoastaan ihmisten peleistä koottu pohjaaineisto Parhaiten testitalanteissa suoriutui BOT1 ja heikoimmin BOT2.	<i>Active learning strategy</i> , jossa järjestelmä kysyy uuden tilanteen edessä ihmiseltä neuvoa ongelman ratkaisemiseen
(Hou ym., 2017)	Evolutiivisen siirtovahvistusoppimisen viitekehysten (<i>evolutionary transfer reinforcement learning framework</i> , eTL) tehokkuuden testaaminen huipputason TL-menetelmiä vastaan MAS:ssa (monitoimija-järjestelmissä)	Ensin viitekehysten toimintaa validoitiin "miinakenttien navigointi" MNT-alustalla ja sen jälkeen sen toimintaa testattiin <i>Unreal Tournament 2004</i> ammuntapelissä	Mahdollisten tilojen määrä on valtava Kaikkien mahdollisten toimintojen ja tilojen määrittely on mahdotonta ja ennakkomäärittelyyn liittyy luotettavuusriski Tekoälyllä on vaikeuksia erottaa hyvät ja huonot neuvot toisistaan <i>blind reliance</i>	eTL suoriutui tehtävissä tavallista PTL:ää paremmin Tutkimusta eTL-strategioiden kehittämisessä kannattaa laajentaa	eTL keskiössä on <i>memetic automation</i> eli viitekehys, jossa tekoäly oppii kulttuurisen informaation itsenäisesti tai vuorovaikutuksessa. Agentilla on valmiiksi olemassa tietyt meemirepresentaatiot, jotka päivittyvät ja muotoutuvat (<i>meme assimilation</i>) evoluution kautta.
(Li ym., 2020)	Epätarkasti määriteltyjen ongelmien ratkomiseen liittyvien haasteiden tarkastelu ja <i>Extenics</i> -teoriaan perustuvien mahdollisten älykkäiden ratkaisumallien analysointi	Analysoidaan syitä siihen, miksi epämääräisiä ongelmia on vaikea ratkaista älykkäästi ja esitetään <i>extenics</i> -teorian soveltamista älykkään tiedon hyödyntämiseksi	Puutteellinen ja piilossa oleva tieto Ongelma ei ole selkeä tai tarkkaan määritelty Hyödyllisen tiedon tuottamiseen ei ole mekanismeja ja kaiken kattavan tiedon tuottaminen on haastavaa	Alustavia tutkimussuuntia tulevaan kehitystyöhön	<i>Extenics</i> -teorian ajatuksena on laajentaa ongelmanratkaisustrategioita, jotta epätarkasti määriteltyjen ongelmien ratkaisu onnistuisi. Ihmisen ja koneen välisen tiedonkäsittelyn yhdistäminen

(Jones & Tanimoto, 2018)	Määritellä ja kehittää prototyyppi tutkimuskeskeisestä konealgebrasta, joka kykenee ratkomaan ammattilaisten määrittelemiä ongelmia	Matemaattisten tutkivien tehtävien suorittamista tukevan prototyypin kehittäminen ja sen toimintaperiaatteiden kuvaaminen	Tekoöly rajoittaa luovuttaa ideointivaiheessa ja tilanteissa, joissa ei ole selkeää tavoitetta	Järjestelmä, joka ei tuota perinteisen tekoölyn tapaan tarkkaan määriteltyä syötettä, vaan sen sijaan etsii hakuavaruudesta useita erilaisia syötteitä, jotka voisivat vastata samaa lauseketta	Tutkiva ongelmanratkaisu, jossa ihminen ja tekoöly työskentelevät yhdessä: ihminen huolehtii intuitiosta ja kehittää jatkuvasti haun suuntaa, kun taas tekoöly hyödyntää nopeuttaan ja muistiin tehäkseen perusteellisia hakuja
(Hwang ym., 2018)	Luokitella autonomisen koneoppimisen tyypilliset ongelmanratkaisun vaiheet, esitellä ongelmanratkaisun prosessi sekä ehdottaa itsenäisen koneoppimisen mallinnusmenetelmää hyödynnettäväksi tehtävien suorituksessa	Autonomisen koneoppimisen toteutus esitellään tehtäväontologian perusteella Tarkastelussa on erityisesti kolme keskeistä näkökulmaa: menetelmä koneoppimisen sanaston hankkimiseksi, tiedon esittäminen käsittegraafina, UML-pohjainen metamalli toteutusta varten	Autonominen koneoppiminen on vielä lapsen kengissä	Avainsanat koneoppimisen ontologian rakentamiseksi <i>Workflow</i> , jonka avulla asiantunteimmat käyttäjät kykenevät suorittamaan monimutkaisia tehtäviä ja voivat helposti toteuttaa koneoppimismallin tiettyyn sovellukseen	Ehdotettu mallinnusmenetelmä voi vähitellen kartuttaa tietoa koneoppimisesta lisäämällä tiettyyn työprosessiin automaattisen koneoppimisen moduulin Automaattisen koneoppimisen tekniikat tuottavat mahdollisuuden turhiin tehtävien vähentämiseen
(Correia 2019)	Joukko-tekoölyn sosio-tekniisen infrastruktuurin keskeisten ominaisuuksien tunnistaminen teoreettisen viitekehysten luomiseksi Tuottaa tutkimuskirjallisuuden pohjalta uusi heuristiikkojen joukko	Kirjallisuuden rakenteellinen jäsentäminen kuvailevan katsauksen kautta: käsittekarttaa ja taksonomian kehity.	Koneoppimisalgoritmit edellyttävät laajaa ja virheetöntä harjoitteluaaineistoa toimiakseen tehokkaasti Tekoölyn ja ihmisen välisen yhteistyön sosioteknisiä järjestelmiä tai hybridi-älykkyyden muotoja ei vielä ole	Ihmiskeskeisen tekoölyn viitekehys, joka sisältää seuraavat näkökulmat: 1) Aika/tila 2) Joukon käyttäytyminen 3) Tehtävät 4) Motivaatiotekijät 5) Laaturolli 6) Tilanne 7) Eettiset näkökulmat ja yksityisyys 8) Hallinnointi 9) Ohjelmointialusta 10) Päätelykyvyt 11) Harjoittelustrategiat 12) Skaalautuvuus 13) Älykkyyden risteys	Joukkoistaminen voisi olla luotettava menetelmä ohjattuun ja aktiiviseen oppimiseen: Päätelykyvyn lisääminen hybridiälykkyyden järjestelmiin tukee parempia päätöksiä, koska kone voi oppia joukon käytöksestä Yhteisevoluutionäkökulma: ihmisen ja koneen kyvyt kehittyvät rinnakkain ajan kanssa
(Mokhtari 2019)	Artikkelissa laajennetaan EBPD viitekehystä lisäämällä jokaiseen toimintakaavioon (<i>activity schema</i>) representaatio sen soveltuvuuden laajuudesta	Toimintakaavion soveltuvuutta testattiin <i>Three-Valued Logic Analysis</i> (TVLA) mallilla TVLA:n avulla arvioitiin myös, voidaanko nykyistä toimintakaaviota hyödyntää tiettyjen ongelmien ratkaisemiseen Työ validoitiin kahdella perinteisellä suunnittelujärjestelmällä ja simuloitiin PR2:lla Gazebo-ympäristössä.	Tavoitteiden saavuttamiseksi edellytettävien toimintojen suunnittelu on vaikeaa Klassisia suunnittelutapoja (tapauspohjainen suunnittelu ja makro-operaattorit) koskevan tiedon oppimisen haasteena on usein varastointi, indeksointi ja tiedon hakeminen	Alustavat tulokset osoittavat hyvää skaalautuvuutta	Experience-based <i>planning domains</i> (EBPDs) eli kokemuspohjaiset suunnittelujärjestelmät. EBPD mallissa ei tarvitse tallentaa laajoja konkreettisten tapauspohjien pohjia, vaan se yhdistää yleistyksen ja abstraktoinnin eri muotoja. Kokemuksien kartuttaminen tapahtuu ihmisen ja koneen välisessä vuorovaikutuksessa.

4.2 Keskeiset ongelmat tekoälyn ongelmanratkaisukyvyssä

Olen Taulukko 4 ryhmitelty aineistosta esiin nousseet keskeiset haasteet tekoälyn ongelmanratkaisukyvyssä. Ryhmittelyn olen laatinut suhteessa muussa kirjallisuudessa tunnistettuihin ongelmiin. Taulukosta on helposti havaittavissa, että kirjallisuushakuni pohjalta löytyneissä artikkeleissa painottuvat frame-ongelma ja tiedon hankinnan ongelma.

Ongelmien luokittelussa on viimeisenä kohtana vielä ns. muut -kategoria, koska kaikki artikkeleissa esiin nostetut ongelmat eivät sopineet teoriaosuuden mukaiseen luokitteluun. Tekniset toteutukset eivät varsinaisesti ole tutkimukseni ensisijainen kiinnostuksen kohde (ks. kappale 1.3), joten mainitsen tässä lyhyesti esiin nousseet kaksi teknisen toteutuksen puutetta. Hwang ym. (2018) mukaan autonominen koneoppiminen, jonka avulla voitaisiin vähentää turhia tehtäviä, on vielä lapsen kengissä. Vastaavasti teknisenä haasteena Correia ym. (2019) mainitsevat, että tekoälyn ja ihmisen väliseen yhteistyöhön ei ole onnistuttu kehittämään sosioteknisiä järjestelmiä tai hybridi-älykkyyden muotoja

Taulukko 4. Aineistosta tunnistetut tekoälyn ongelmat

Ongelma	Artikkelin näkökulma	Artikkelit, joissa ongelma on mainittu
Frame-ongelma	Ongelma tai tavoite ei ole selkeä Mahdollisten tilojen valtava määrä	(Li ym., 2020) (Jones & Tanimoto, 2018) (Buligi ym., 2020) (Hou ym., 2017)
Assosiaatiofunktion ongelma		
Symbolisten merkitysten ongelma		
Tiedon hankinnan ongelma	Mallintaminen on työlästä tai mahdotonta Aineiston luotettavuusriski	(Moral ym., 2020) (Li ym., 2020) (Mokhtari ym., 2019) (Jones & Tanimoto, 2018) (Hou ym., 2017) (Correia ym., 2019) (Buligi ym., 2020)
Muut	Tekninen toteutus puuttuu	(Hwang ym., 2018) (Correia ym., 2019)

4.2.1 Frame-ongelma

Kuten kappaleessa 4.2.1 esittelin, frame-ongelmassa on kyse haasteesta tunnistaa ongelman ratkaisemisen kannalta oleellisia asioita sekä arvioida miten tietty toiminta muuttaa kohteiden ominaisuuksia. Kirjallisuuskatsauksen artikkeleissa tähän liittyviä näkökulmia toivat esiin Li ym. (2020) sekä Jones ja Tanimoto (2018).

Li ym. (2020) puhuvat ongelmien määrittelyn vaikeudesta: yritykset ja organisaatiot kohtaavat sisäisiä ja ulkoisia ympäristöjä, joissa ongelmat ovat monimutkaisia, epämääräisiä ja moniulotteisia. Tavoite voi olla selkeä, mutta ympäröivät olosuhteet sisältävät useita ja muuttuvia epävarmuustekijöitä. Tällaisten ongelmien ratkaisu koneoppimisen tai tiedon päättelyn menetelmin on Li ym. (2020) mielestä haastavaa, mutta ihmiset toivoisivat silti saavansa tekoälyltä kyllä ja ei vastauksia laajempaa tukea ja jopa automaattisia tai semi-automaattisia ratkaisuja. (Li ym., 2018.) Vaikka Li ym. (2020) eivät mainitse frame-ongelmaa, vastaa yllä oleva kuvaus ongelmanratkaisun haasteista mielestäni täysin frame-ongelman mukaista haastetta käytännön tilanteissa vallitsevasta muuttujien monimutkaisuudesta.

Jones ja Tanimoto (2018) lähestyvät aihetta tavoitetilan määrittelyn vaikeuden näkökulmasta, mutta mielestäni tämäkin voidaan tulkita frame-ongelmaan liittyväksi haasteeksi. Jones ja Tanimoto (2018) tuovat esiin tilanteet, joissa tutkijat ja tekoäly työskentelevät yhdessä jonkin tavoitteen eteen, jota ei ole tarkkaan määriteltä tai se ei ole helposti ilmaistavissa. Perinteinen tekoäly vaatii tarkkaan määritellyn tavoitetilan tai optimointifunktion, mutta tutkivassa ongelmanratkaisussa tällaista ei ole mahdollista määrittellä. Mielestäni tässäkin tapauksessa kyse on ongelmanratkaisuun vaikuttavien muuttujien määrittelyn vaikeudesta.

Ongelmanratkaisun monimutkaisen ympäristön lisäksi kirjallisuuskatsauksen artikkeleista oli tunnistettavissa frame-ongelmaan liittyvä haaste kaikkien mahdollisten toimintojen ja tilojen määrittelyn vaikeudesta, minkä toivat tulkintani perusteella esille Buligiu ym. (2020), Hou ym. (2017) ja Mokhtari ym. (2019).

Hou ym. (2017) käsittelevät kaikkien mahdollisten toimintojen ja tilojen määrittelyn haasteetta monitoimijajärjestelmän¹² (*Multiagent systems*, MAS) yhteydessä. Yleensä jokaisen agentin toiminta on ennakkoon määriteltä suhteessa ympäristössä mahdollisesti kohdattaviin tiloihin. Tämä toteutus voi kuitenkin olla epäluotettava, kun ympäristö muuttuu monimutkaiseksi ja muuttuu ajan myötä. Tällaisessa tilanteessa mahdollisten tilojen avaruus on suunnaton. Myös Buligiu ym. (2020) käsittelevät artikkelissään tila-avaruuskaavioita, joilla on valtava koko. Yhtenä esimerkkinä tästä ovat epädeterministiset polynomiset¹³ (*nondeterministic polynomial*) ongelmat, kuten Blocks world¹⁴, joiden tehokkaaseen ratkaisemiseen ei tiedetä olevan polynomisiin algoritmeihin pohjautuvia ratkaisuja. (Buligiu ym., 2020). Myös Mokhtari ym. (2019) mainitsevat tila-avaruuskaavioihin liittyvän polynomisen ajan haasteen ja valitsevat siksi täydellisen suorituksen sijaan algoritmin, joka suoriutuu polynomisessa ajassa ja on siten tehokkaampi.

Edellä kuvattujen esimerkkien pohjalta voi helposti päätellä, että frame-ongelmaan sisältyvä haaste tehtäväympäristön ominaisuuksien määrittelystä

¹² Monitoimijajärjestelmä tarkoittaa sellaista järjestelmää, jossa useampi itsenäinen agentti toimii samanaikaisesti saman tai eri tavoitteiden parissa

¹³ Tieteen termipankin (2023) mukaan epädeterministinen tarkoittaa sellaista algoritmia, joka ei pysty joka askeleellaan valitsemaan oikeaa jatkovaihtoehtoa

¹⁴ Blocks world on tekoälytutkimuksessa paljon käytetty ohjelma, jossa tekoälyn tehtävänä on rakentaa pöydällä olevista palikoista tietyntylaisia pinoja siirtäen kerralla vain yhtä palikkaa.

muuttuu sitä vaikeammaksi, mitä monimutkaisemmasta tilanteesta on kyse. Samalla myös tila-avaruuskaavioiden koko kasvaa ja voi käytännössä olla jopa ääretön, esimerkiksi Buligiū ym. (2020) laskelmien mukaan Blocks world ongelmassa mahdollisten siirtojen määrä on 30 palikan tilanteessa yli 5^{36} .

4.2.2 Tiedon hankinnan ongelma

Li ym. (2020) käsittelevät tiedon hankinnan ongelmaa artikkelissaan kattavasti, vaikka eivät käytäkään termiä *knowledge acquisition*. Artikkelissa nostetaan esiin tiedon hankinnan ongelman kannalta tyypillisiä näkökulmia ja esitellään miksi epämääräisiä ongelmia on vaikea ratkoa älykkäästi: datan luokittelu on haastavaa, datasta on haastava hankkia oleellinen tieto, osa tiedosta on hiljaista tietoa ja piilossa, tietoa ei ole riittävästi, olosuhteiden ja tavoitteiden välillä on tietokuilu, ongelmanratkaisusta ja laskennasta puuttuu oleellista tietoa jne. (Lu ym., 2018.)

Tiedon hankintaan liittyvät ongelmat olivat selkeästi vahvimmin esillä tutkimusaineistossa tunnistetuista tekoälyn ongelmanratkaisuun liittyvistä haasteista. Oman pohdintani perusteella tämä voisi johtua muun muassa käsitteen laajuudesta; sen alle on helppo yhdistää useita näkökulmia. Toisaalta tieto ja sen käsittely on yksi ongelmanratkaisun keskeisiä osia, joten tiedon hankinnan ongelman ratkaisu on kriittistä älykkyyden kehittämisen näkökulmasta. Artikkeleissa tunnistetut ongelmat voidaan mielestäni jakaa tiedon hankinnan ongelman osalta kolmeen kategoriaan: 1) kattavan koulutusaineiston tai tietopohjan rakentaminen on työlästä tai mahdotonta 2) aineiston luotettavuus ja mahdolliset vinoumat 3) musta laatikko -ongelma.

Moral ym. (2020) käsittelevät artikkelissaan ongelmanratkaisua Etelä-Amerikkalaisen Truco korttipelin näkökulmasta. Heidän mukaansa monimutkaiseen peliympäristöön on uuvuttavaa mallintaa kaikki mahdolliset ongelmat ja ratkaisut, jotka agentti voi pelitilanteessa kohdata. Samoilla linjoilla ovat myös Mokhtari ym. (2019), joiden mukaan tavoitteiden saavuttamiseksi edellytettävien toimintojen suunnittelu on vaikeaa ja edellyttää usein tasapainoilua tarvittavien tietojen ja laskentatehon välillä. Mallintamisen haasteeseen liittyvät mielestäni myös Buligiū ym. (2020) mainitsevat ongelman määrittelyssä olevat *deadlocks*¹⁵, jotka vaikuttavat siihen, onko ongelma ratkaistavissa polynomisessa ajassa¹⁶.

Tiedon hankinnan ongelman näkökulmaksi tulkitsen myös Moral ym. (2020) pohdinnat siitä, miten informaatio saadaan muutettua hyödynnettävään muotoon sekä miten oleellinen tieto saadaan rakennettua järjestelmään, siten että sitä pystytään mielekkäällä tavalla käsittelemään. Tähän viittaa myös Mokhtari ym. (2019) maininta siitä, että kokemuspohjaisten suunnittelujärjestelmien tiedon oppimisen haasteena ovat usein varastointi, indeksointi ja tiedon hakeminen.

¹⁵ Deadlockseilla tarkoitetaan tilannetta, jossa algoritmi ei pysty etenemään, koska yksi tai useampi prosessi odottaa tulevaa syötettä.

¹⁶ Polynomisen aika viittaa sellaiseen algoritmiin, joka on laskettavissa syötteen kokoon nähden kohtuullisessa ajassa. Suuremmalla syötteellä ongelma ei olisi enää laskettavissa polynomisessa ajassa.

Li ym. (2020) näkemys asiaan on, että hyödyllisen tiedon tuottamiseen ei ole mekanismeja ja kaiken kattavan tiedon tuottaminen on haastavaa.

Tiedon hankinnan ongelmaan liittyvät kiinteästi myös nykyisten koneoppimisen menetelmien osalta tunnistetut ongelmat, jotka koskevat luotettavan aineiston tuottamista. Tiedon käsittelyn kannalta oleelliset tiedot voivat olla puutteellisia, virheellisiä tai vinoutuneita. Vastaavia huomioita nostettiin esiin useissa artikkeleissa, joten luotettavaa aineistoa koskevan haasteen voisi tulkita olevan varsin yleinen ongelma. Vinouman vaaroista mainitsevat muun muassa Moral ym. (2020), joiden mukaan tapauskohtaiseen päättelyyn liittyy vinouman riski: vastaavaa esimerkkitalannetta ei ole tai tietty toiminta on yliedustettuna opetusmateriaalissa. Opetusmateriaalin tai kouluttajan luotettavuuden riskistä puhuvat myös Hou ym. (2017), joiden mukaan tekoälyllä on vaikeuksia erottaa hyvät ja huonot "neuvot" toisistaan. Tätä kutsutaan sokean riippuvuuden (*blind reliance*) ongelmaksi. Correia ym. (2019) ovatkin sitä mieltä, että koneoppimisalgoritmit edellyttävät laajaa ja virheetöntä harjoitteluaineistoa toimiakseen tehokkaasti.

Jones ja Tanimoto (2018) eivät suoraan mainitse mustan laatikon ongelmaa, mutta nostavat esiin saman teeman eli sen, että tekoälyn valitsevat ratkaisut eivät ole läpinäkyviä. Heidän artikkelissaan tämä nostetaan esiin tutkijoiden ja tekoälyn välisen tutkivan työskentelyn osalta: professorit turvautuvat paperiin ja kynään, koska tekoäly rajoittaa heidän luovuuttaan ja sen valitsevat etenemismallit eivät ole läpinäkyviä.

4.3 Ratkaisuja tekoälyn ongelmanratkaisukyvyyn kehittämiseksi

Esittelen tässä kappaleessa kirjallisuuskatsauksesta nousseita ratkaisuehdotuksia tekoälyn ongelmanratkaisukyvyyn kehittämiseksi. En käy jokaista ratkaisuehdotusta yksityiskohtaisesti läpi, vaan tarkastelen niitä suhteessa tunnistettuihin ongelmiin. Lisäksi keskityn erityisesti niihin artikkeleihin, joiden ratkaisuehdotukset vastaavat mielestäni parhaiten tutkimuskysymykseeni ja vaikuttavat omasta mielestäni kehityskelpoisilta. Olen ryhmitellyt ratkaisuehdotukset Taulukko 5 mukaisesti suhteessa aiemmin tunnistettuihin ongelmiin. Vaikka ratkaisuehdotukset olivat kaikissa artikkeleissa erilaisia, muutama yhdistävä tekijä oli tunnistettavissa. Selkeästi yleisin (viisi artikkelia kahdeksasta) ehdotus oli, että ihmistä kannattaisi hyödyntää ongelmanratkaisuprosessissa kouluttajana, tiedon validoijana tai tekoälyn kumppanina.

Menetelmällisesti artikkelit painottuivat koneoppimisen menetelmiin tai koneoppimisen pohjalta kehitettyihin uusiin menetelmiin. Tarkastelussa olivat automaattinen koneoppiminen (Hwang ym., 2018), ohjattu oppiminen (Correia ym., 2019), siirto-oppiminen (Hou ym., 2017), aktiivinen oppiminen (Moral ym., 2020), kokemuspohjainen oppiminen (Mokhtari ym., 2019) sekä extenics-teoria (Li ym., 2020). Ongelmanratkaisumenetelmänä mainittiin tila-avaruushaku (Buligiu ym., 2020; Hou ym., 2017), tehtäväontologia (Hwang ym., 2018), tapauskohtainen päättely (Moral ym., 2020), kokemuspohjaiset suunnittelujärjestelmät (Mokhtari ym., 2019), tietokonealgebra (Jones & Tanimoto, 2018) ja joukkoistaminen (Correia ym., 2019).

Seuraavissa kappaleissa esittelen tarkemmin artikkeleissa esitettyjä ratkaisuehdotuksia ja tarkastelen niiden puutteita ja mahdollisuuksia tekoälyn monimutkaisen ongelmanratkaisukyvyyn vahvistamiseksi. Kappaleet olen otsikoinut artikkeleista tunnistamani ongelmien mukaisesti, mutta ratkaisunäkökulmasta. Olen pyrkinyt tuomaan jokaisen kappaleen yhteydessä kustakin ratkaisuehdotuksesta esille sellaiset näkökulmat, jotka vastaisivat nimenomaiseen ongelmaan. Mikäli jokin ratkaisuehdotus vastaa useampaan ongelmanäkökulmaan, olen esitellyt sitä useammassa kappaleessa.

Taulukko 5. Aineistosta tunnistetut ratkaisuehdotukset

Ongelma	Artikkelin ratkaisuehdotus	Artikkelit, joissa ongelma on mainittu
Frame-ongelma Ongelma tai tavoite ei ole selkeä	Extenics-teoria Ihmisen ja tekoälyn yhteistyö	(Li ym., 2020) (Jones & Tanimoto, 2018) (Li ym., 2020)
Frame-ongelma Mahdollisten tilojen valtava määrä	Tila-avaruuskaavioiden koon mittaaminen Evolutiivinen siirtovahvistusoppimisen viitekehys	(Buligiu ym., 2020) (Hou ym., 2017)
Tiedon hankinnan ongelma Tiedon mallintaminen ontyölästä tai mahdotonta	Tapauspohjaisen oppimisen strategiat Kokempuspohjaiset suunnittelujärjestelmät (EBPD) Autonominen koneoppiminen	(Moral ym., 2020) (Mokhtari ym., 2019) (Hwang ym., 2018)
Tiedon hankinnan ongelma Aineiston luotettavuusriski	Evolutiivinen siirtovahvistusoppimisen viitekehys Joukkoälykykyys	(Hou ym., 2017) (Correia ym., 2019)

4.3.1 Ongelman selkiyttäminen

Kuten kappaleessa 2.3 todettiin, monimutkaiselle ongelmalle on tyypillistä, että tavoite ja käytettävissä olevat tiedot eivät ole selkeitä. Tekoälyllä on tällaisissa tilanteissa vaikeuksia määritellä tai hahmottaa toivottu tavoitetila, koska ympäristössä ja olosuhteissa on paljon epävarmuustekijöitä. Jotta epätarkasti määriteltyjen ongelmien ratkaisu onnistuisi Li ym. (2020) ovat tutkineet ja kehittäneet Extenics-teoriaa, jonka ajatuksena on laajentaa ongelmanratkaisustrategioita. Kirjallisuushakuni tuottamassa artikkelissa Li ym. (2020) ehdottavat Extenics-teorian yhdistämistä tekoälyn ja tiedon johtamisen kanssa, jotta epätarkasti määriteltyjen ongelmien ratkaisu onnistuisi. Artikkelin mukaan Extenics pohjautuu professori Yangin kehittämään älykkään tietoinformaatiojärjestelmän (*extension information-knowledge intelligence formal system*) laajennuksen viitekehukseen. Tämän tarkemmin artikkelissa ei kuvata, mistä teoriassa on käytännössä kyse. Kokonaiskuvan hahmottamiseksi selvitin toisesta Li ym. (2022) artikkelista, että Extenics -teoria pohjaa logiikkaan ja matematiikkaan. Teorian tarkoituksena on

asioiden, objektien, suhteiden, informaation ja tiedon muuttaminen esitettäväksi loogiseen muotoon. (Li ym., 2022.) Laajemman ymmärryksen muodostaminen teorian käyttökelpoisuudesta edellyttäisi tämän tutkimuksen mittasuhteet ylittävää perehtymistä aiheeseen. Tiiviin kuvauksen ja tulkinnan pohjalta Extenics vaikuttaisi kiinnostavalta teorialta nimenomaan Frame-ongelmaan liittyvien haasteiden ratkaisemisen näkökulmasta. Mikäli tekoälylle olisi mahdollista Extenics-teorian pohjalta luoda looginen viitekehys asioiden, objektien, suhteiden ja tiedon välisistä ominaisuuksista, uskoisin sen edistävän merkittävästi myös tekoälyn kykyä monimutkaiseen ongelmanratkaisuun. Koska Li ym. (2020) esittävät artikkelissaan kuitenkin ihmisen ja koneen välisen tiedonkäsittelyn yhdistämistä ja molempien vahvuuksien hyödyntämistä, epäilen että Extenicsin mukaisen viitekehysten hyödyntäminen tekoälyn ongelmanratkaisussa ei vielä ole kovin pitkällä. Artikkelissa ei esitetä asiasta konkreettista näyttöä tai tutkimusasetelmaan perustuvia tuloksia, joten todellisen sovellettavuuden arvioiminen on haastavaa.

Myös Jones ja Tanimoto (2018) päätyivät artikkelissaan ihmisen ja koneen vahvuuksia yhdistävään ratkaisuehdotukseen tutkivan ongelmanratkaisun näkökulmasta. Heidän mukaansa paras lopputulos saadaan, kun ihminen ja tekoäly työskentelevät yhdessä siten, että ihminen huolehtii intuitiosta ja kehittää jatkuvasti haun suuntaa, kun taas tekoäly hyödyntää nopeuttaan ja muistiaan tehdäkseen perusteellisia hakuja. Jones ja Tanimoto (2018) tutkimuksen tuloksena oli järjestelmä, joka ei tuota perinteisen tekoälyn tapaan tarkkaan määritettyä syötettä, vaan sen sijaan etsii hakuvaruudesta useita erilaisia syötteitä, jotka voisivat vastata samaa lauseketta. Ihmisen tehtäväksi jää arvioida, mikä vaihtoehtoista on jatkotyöskentelyn kannalta sopivin. (Jones & Tanimoto, 2018.)

Molemmissa edellä kuvatuissa artikkeleissa lopputuloksena oli ajatus ihmisen ja koneen välisestä yhteistyöstä ja molempien vahvuuksien hyödyntämisestä. Tulkitsen tämän kuvastavan sitä, että Frame-ongelman ratkaiseminen on haastavaa. Yksikään kirjallisuuskatsauksen artikkeleista ei mielestäni tarjonnut varsinaista ratkaisua siihen, miten tekoälyn kykyä oleellisuusarvioiden tekemiseen ja toiminnan muutosten vaikutusten tunnistamiseen voitaisiin kehittää.

4.3.2 Mahdollisten tilojen määrän hallinta

Toisena Frame-ongelmaan liittyvänä näkökulmana artikkeleista nousi esiin mahdoton tehtävä määrittellä ennakkoon kaikki monimutkaisen ja muuttuvan tilanteen edellyttämät säännöt. Tehtävän haastavuuden lisäksi siihen liittyy myös epäluotettavuuden riski, sillä monimutkaisessa ongelmassa muuttujat ja tilanteet voivat myös vaihdella.

Buligiu ym. (2020) esitelevät ja kokeellisesti arvioivat laskennallisia algoritmeja, joilla voidaan mitata monimutkaisten ongelmien hakukaavioita. Tutkimuksen kohteena on Blocks world¹⁴ ongelma, jonka osalta Buligiu ym. (2020) ehdottavat uusia mittareita ongelman koon ja monimutkaisuuden määrittämiseksi, esittävät algoritmeja mainittujen mittareiden arviointiin sekä arvioivat kokeellisesti ehdotettuja algoritmeja. Lopputuloksena Bugliu ym. toteavat, että tutkimuksessa esitetty algoritmi on tehokas hakukaavioiden mittaamiseen. he myös pohtivat, että analyysin tuloksia ja ehdotettuja algoritmeja voitaisiin laajentaa muiden mittareiden (mm. deadlocks kokoonpanojen määrä) sekä tilojen

monimutkaisuuden arviointiin myös muissa ongelmissa, kuten esimerkiksi BW4T:ssä (Blocks world tiimityöskentelynä). (Buligiu ym., 2020.)

Omasta näkökulmastani tarkasteltuna Buligiu ym. (2020) ehdotus on melko tekninen eikä varsinaisesti tarjoa ratkaisua siihen, miten toiminnan vaikutusten määrittämisen haaste saataisiin ratkaistua. Pohdin kuitenkin, voisiko monimutkaisten ongelmien hakukaavioiden laajuuden mittaaminen auttaa hahmottamaan ongelmaympäristön laajuutta sekä ymmärtämään, miten tekoäly on toiminut ongelmanratkaisussa erilaisissa tilanteissa. Hieman tämän suuntaista ajatusta hyödynsivät omassa tutkimuksessaan Mokhtari ym. (2019), jotka hyödynsivät kolmearvoisen loogisen rakenteen (*Three-Valued Logic Analysis TVLA*) työkalua esittämänsä ratkaisun soveltuvuuden testaamiseen. TVLA:n avulla laskettiin toimintakaavioiden sovellettavuutta ja arvioitiin, voidaanko nykyistä toimintakaaviota hyödyntää tiettyjen ongelmien ratkaisemiseen. Tämä ei kuitenkaan ollut Mokhtari ym. (2019) varsinainen ratkaisuehdotus tekoälyn ongelmanratkaisun kehittämiseen, vaan ratkaisuksi esitetään kokempohjaisten suunnittelujärjestelmien kehittämistä, mitä esittelen tarkemmin kappaleessa 4.3.3.

Hou ym. (2017) esittämä ratkaisuehdotus evolutiivisen siirtovahvistusoppimisen mallista (*evolutionary transfer reinforcement learning framework* eli eTL) kytkeytyy mielestäni sekä frame-ongelmaan, että tiedon hankinnan ongelmaan. Frame-ongelman näkökulmasta malli tarkastelee älykkään agentin toimintaa monitoimijajärjestelmässä (MAS), johon tarvitaan erilaisiin tilanteisiin sopeutuva agentti. Monitoimijajärjestelmässä agenttien välinen tiedonvaihto mahdollistaa siirto-oppimisen (*transfer learning*), jossa hyödynnetään aikaisempia ja hyvin tutkittuja ongelmatilanteita käsillä olevan tilanteen ongelmanratkaisumahdollisuuksien parantamiseksi. Agenttien välisen tiedon vaihdon haasteena on kuitenkin sokea riippuvuus eli vaikeus tunnistaa, mitkä lähdetehtävistä ovat hyödyllisimpiä: paras opettaja ei välttämättä olekaan sillä hetkellä parhaiten suoriutuva agentti, vaan oleellista on tilanteen ja tiedon samankaltaisuus. Tähän ongelmaan Hou ym. (2017) esittävät ratkaisuksi memeettisen evoluution viitekehystä, jonka tarkoituksena on luoda suurempaa joustavuutta kyvyssä ratkaista monimutkaisia ongelmia vahvistamalla kehitystä ja vaihtelua agenttien maailmaa koskevan tiedon osalta. Tutkimuksessa eTL osoitti meemien valikoinnin (*meme selection*) vähentävän sokeaa riippuvuutta. (Hou ym., 2017.) Evolutiivisen siirtovahvistusoppimisen tiedon hankinnan periaatteita esittelen myös kappaleessa 4.3.4.

4.3.3 Tiedon mallintaminen

Moral ym. (2020) tarkastelevat ongelmanratkaisua Truco-korttipelin näkökulmasta, joka on suosittu Etelä-Amerikassa. Trucossa onnella, strategialla ja psykologialla on pokerin tapaan merkittävä rooli, mikä tekee siitä omalla tavalla monimutkaisen pelin. Pelissä on myös mahdollista bluffata. Trucolle ei ole olemassa valmiiksi mallinnettua opetusaineistoa, jonka pohjalta tekoälyä voisi kehittää, joten Moral ym. (2020) ovat selvittäneet tapauskohtaisen päättelyn ja aktiivisen oppimisen yhdistämistä peliagentin tarvitseman tiedon hankkimiseksi.

Tietokonepelien agenttien odotetaan toimivan ihmisen kaltaisesti ongelmanratkaisutilanteissa. Tällaisissa tilanteissa tapauskohtaisen päättelyn (CBR) pohjalta toimivat järjestelmät oppivat ja ratkaisevat ongelmia hyödyntämällä

aikaisemmin tallennettuja tapauksia tai esimerkkejä. Ajatuksena on, että keskenään samankaltaisissa tilanteissa voidaan soveltaa aiemmin hyödynnettyjä ratkaisuja. Tämän mallintaminen on kuitenkin työlästä, joten Moral ym. (2020) mielestä on tarpeen kehittää tekniikoita, jotka mahdollistavat agenttien oppimisen, vaikka alkuperäinen tietämys ongelmanratkaisusta olisi rajoitettua.

Moral ym. (2020) tarkastelevat artikkelissaan erilaisia tiedon oppimisen tekniikoita viiden BOT:n välisissä pelitilanteissa. Agentit suoriutuivat testattavana olleiden uusien oppimistekniikoiden avulla pelistä paremmin, kuin jos käytössä oli ainoastaan aiempiin peleihin pohjautuva tapausvalikko (513 ihmisen pelaamaa peliä). (Moral ym. 2020.) Tutkimuksen keskeisenä tuloksena Moral ym. (2020) totesivat, että hyödyntämällä erilaisia tapauspohjaisen oppimisen strategioita, agentit suoriutuivat pelistä paremmin, kuin pelkällä tapauskohtaisella päättelyllä. Oppimiseen perustuvien menetelmien etuna oli, että järjestelmä ei toista aina rajallista määrää samanlaisena toistuvia ratkaisuja, vaan oppii myös uutta. (Moral ym. 2020.) Tutkimuksessa esitetyt näkökulmat tiedon hankinnan laajentamisesta tapauskohtaisen päättelyyn pohjautuvien menetelmien osalta vaikuttivat mielestäni varteenotettavilta tavoilta ehdotuksilta tekoälyn ongelmanratkaisukyvyyn kehittämiseksi.

Tiedon hankinnan haasteita ja mahdollisia ratkaisuja tarkastelevat myös Mokhtari ym. (2019), jotka ovat tutkineet kokempohjaisia suunnittelujärjestelmiä (EBPD) viitekehyksenä pitkäaikaisen oppimisen ja suunnittelun integroimiseksi robotiikkaan. Kuten Moral ym. (2020) tapauspohjaisten oppimisten strategioissa, EBPD mallissa ei tarvitse tallentaa laajoja konkreettisten tapausten pohjia, vaan se yhdistää yleistyksen ja abstraktoinnin eri muotoja. Omassa tutkimuksessaan Mokhtari ym. (2019) laajentavat EBPD mallia vielä lisäämällä jokaiseen toimintamalliin representaation sovellettavuuden laajuudesta. Mallin toimintakykyä testattiin tosielämää simuloivalla PR2 robotilla Gazebo -ympäristössä. Tulos oli, että EBPD järjestelmä löytää pääteltyä laajuutta hyödyntäen automaattisesti soveltuvan toimintamallin tehtävänä olevan ongelman ratkaisemiseen.

Mokhtari ym. (2019) ratkaisuehdotus on kiinnostava. Artikkelin haasteena tosin on, että johtopäätökset on kuvattu hyvin suppeasti ja kahvin tarjoilemiseen liittyvä koetilanne ei varsinaisesti ole mielestäni monimutkainen ongelma. Tästä huolimatta EBPD ja sen laajennus representaatioiden sovellettavuuden laajuudesta vaikuttavat toimivalta ja perinteisiä malleja yksinkertaisemmalta tavalla tiedon hankkimiseksi.

Kolmannen ratkaisuehdotuksen tiedon mallintamisen näkökulmaan esittivät tulkintani mukaan Hwang ym. (2018), jotka tarkastelevat autonomisen koneoppimisen mallia. Artikkelissa laaditaan autonomisen koneoppimisen ontologia sekä ehdotetaan itsenäisen koneoppimisen mallinnusmenetelmää hyödynnettäväksi tehtävien suorituksessa. Ehdotetun mallin tarkoitus on autonomista koneoppimista hyödyntäen oppia uudelleen olemassa olevia resursseja ja tuottaa siten uutta ymmärrystä. Artikkelin pääasiallinen sisältö jää mielestäni ohueksi ja tutkijat itsekin toteavat tunnistaneensa keskeisiä käsitteitä ja määritelleensä asioita käsitteellisellä tasolla. Kaikissa kolmessa tässä kappaleessa esitellyssä artikkelissa näkökulmana oli tiedon hankinnan laajentaminen siten, että olemassa olevasta tiedosta opitaan uutta, sitä käsitellään uudella tavalla tai sen tueksi hankitaan täydentävää tietoa.

4.3.4 Tiedon luotettavuuden vahvistaminen

Tiedon luotettavuuden näkökulmasta ratkaisuja tekoälyn kehittämiseksi tarkastelivat sekä Correia ym. (2019) että Hou ym. (2017). Correia ym. (2019) näkökulma on joukkoälykkyyden hyödyntämisessä, kun taas Hou ym. (2017) näkökulmana on meemirepresentaatioiden ja evolutiivisen siirtovahvistusoppimisen viitekehys. Näistä kahdesta mallista Hou ym. (2017) ehdotus on mielestäni tekoälykehityksen näkökulmasta edistyneempi. Artikkelin teemoja esittelin jo kappaleessa 4.3.24.3.2 monitoimijajärjestelmässä toimivien agenttien välisen tiedon siirron osalta, joten keskityn tässä artikkelin näkökulmiin luotettavan tiedon hankinnasta.

Hou ym. (2017) kehittämä evolutiivinen siirtovahvistusoppimisen malli perustuu rinnakkaisen siirto-oppimisen (*parallel transfer learning* PTL) malliin. Rinnakkaisen siirto-oppimisen ongelmana on, että agentit tekevät vääriä johtopäätöksiä ja tulkintoja. Tämän aineiston tulkintaan liittyvän haasteen ratkaisemiseksi Hou ym. (2017) esittävät meemirepresentaatioita ja meemievoluutiota. Tekoälyn yhteydessä meemirepresentaatiot tarkoittavat Hou ym. (2017) mukaan eri ongelmanratkaisutilanteisiin siirrettävissä olevia muunnosmatriiseja, joiden avulla eri ongelma-alueilla voidaan saada parempia hakutuloksia. Evolutiivisen siirtovahvistusoppimisen (eTL) perusajatuksena on, että agentilla on valmiiksi olemassa tietyt meemirepresentaatiot, jotka päivittyvät ja muotoutuvat evoluution kautta. Käytännössä eTL on fuusio *imitate-from-elitist* ja *like-attracts-like* toimintaperiaatteista. Hou ym. (2017) toteuttamissa koetilanteissa eTL -agentti suoriutui tavallista PTL mallia paremmin. (Hou ym. 2017.)

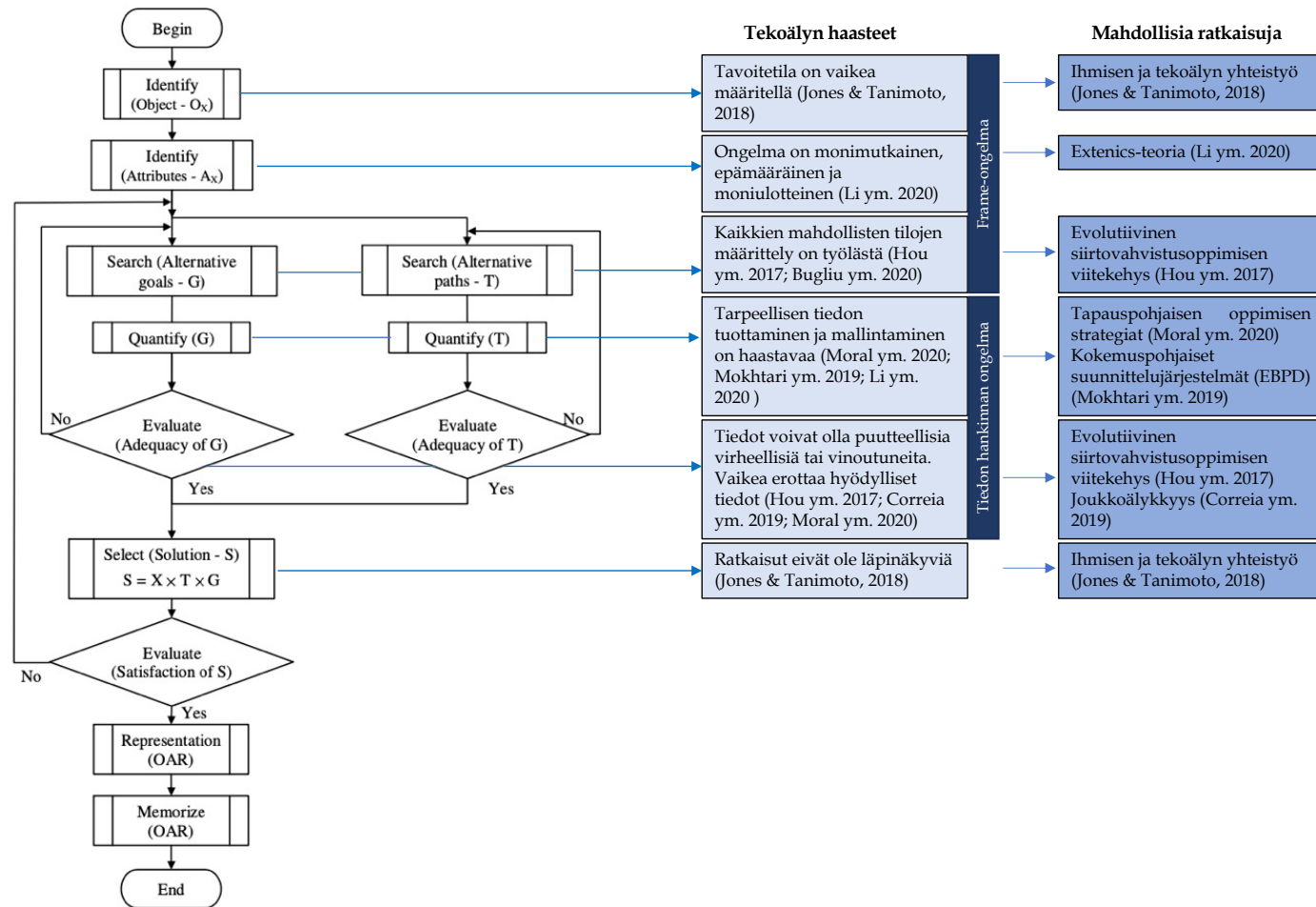
Correia ym. (2019) lähestyvät aineiston luotettavuuden teemaa sosioteknisestä näkökulmasta eli tavoitteena on, että ihmisen ja koneen symbioottisen suhteen kehittäminen ja parempi ymmärrys vie kohti älykkäämpiä järjestelmiä. Correia ym. (2019) ovat toteuttaneet kuvailevan kirjallisuuskatsauksen, jonka tarkoituksena oli tunnistaa joukko-tekoälyn sosio-tekniikan infrastruktuurin keskeisiä ominaisuuksia teoreettisen viitekehityksen luomiseksi. Joukko-tekoälyn ajatuksena on ihmisjoukkojen ja konealgoritmien vahvuuksien yhdistäminen. Käytännössä joukkoälykkyyttä voidaan hyödyntää kaksisuuntaisesti; ihmisen tukena tai ihminen automaation kehittäjänä. Tekoäly voi tehdä joukosta tehokkaamman, taitavamman sekä tarkemman ja vastaavasti joukkoälykkyyden avulla voidaan myös kehittää, ohjata ja täydentää automaatiota. Aineiston luotettavuuden näkökulmasta Correia ym. (2019) mukaan joukkoälykkyyks on tehokas tapa tekoälyn kouluttamiseen ja valvontaan. Päätelykykyjen lisääminen hybridiälykkyyden järjestelmiin tukee parempien päätösten tekemistä myös monimutkaisissa tilanteissa, koska kone voi oppia joukon käytöksestä. Vaikka sosio-tekniikan järjestelmien tutkimus on ollut viime vuosien kiinnostuksen kohteena, toimivaa hybridiälykkyyden yhteisevoluutionäkökulmaa¹⁷ ei ole vielä onnistuttu kehittämään. (Correia ym. 2019.)

¹⁷ Yhteisevoluutionäkökulma tarkoittaa tässä yhteydessä sitä, että ihmisen ja koneen kyvyt kehittyvät rinnakkain ajan kanssa.

5 JOHTOPÄÄTÖKSET

Tässä tutkimuksessa tavoitteenani oli tunnistaa tekoälyn ongelmanratkaisussa ilmeneviä haasteita sekä kartoittaa mahdollisia ratkaisuehdotuksia. Tutkimusote on ollut kognitiivinen eli olen tarkastellut tekoälyn ongelmanratkaisukykyjä suhteessa ihmisen kykyihin. Tutkimusaineistona on ollut systemaattisen kirjallisuuskatsauksen menetelmällä tunnistetut kahdeksan tieteellistä artikkelia, joita olen tarkastellut kappaleessa 1.1 kuvattujen tutkimuskysymysten näkökulmasta.

Kuvioon 9 olen muotoillut oman tulkintani siitä, miten aineistosta nousseet ongelmat tekoälyn ongelmanratkaisussa suhteutuvat ihmisen ongelmanratkaisun kognitiiviseen prosessiin ja minkälaisia ratkaisuehdotuksia niihin on aineistossa esitetty. Kuvion perustana on hyödynnetty Wangin ja Chiewin (2010) mallia ongelmanratkaisun kognitiivisesta prosessista, johon olen suuntaa antavasti yhdistänyt aineistosta nousseet tekoälyn ongelmanratkaisun haasteet ja ratkaisut. Yksittäisen haasteen ja kognitiivisen prosessin vaiheen vastaavuus ei ole täsmällinen vaan enemmänkin viitteellinen. Kuviossa olevat nuolet osoittavat tätä viitteellistä yhteyttä kognitiivisen prosessin ja tunnistettujen ongelmien välillä. Kuvion viimeiseen sarakkeeseen on koottu tutkimusaineistossa esitetyt ratkaisuehdotukset pois lukien Hwang ym. (2018), joiden ratkaisuehdotus ei tulkintani mukaan vastaa mihinkään tunnistetuista haasteista. Kuvio toimii yhteenvetona tutkimukseni keskeisistä tuloksista ja yhdistää tulokset laajempaan teoreettiseen viitekehykseen. Seuraavissa kappaleissa käyn lyhyesti läpi tutkimukseni keskeiset tulokset sekä pohdin tutkimustulosten hyödynnettävyyttä ja jatkotutkimusmahdollisuuksia.



Kuvio 8. Tutkimusaineiston ja teorian pohjalta laadittu tulkinta tekoölyn haasteista ja mahdollisista ratkaisuista kohti ihmisen kaltaista ongelmanratkaisua.

5.1 Keskeisiä ongelmia tekoälyn ongelmanratkaisukyvyssä

Ensimmäisen tutkimuskysymykseni oli: *minkälaisia keskeisiä ongelmia tekoälyn ongelmanratkaisukyvyyn osalta on kirjallisuudessa tunnistettu suhteessa ihmisen ongelmanratkaisukykyyn?* Kuviosta 9 on havaittavissa, että tutkimusaineistosta nousi esiin tekoälyn ongelmanratkaisuun liittyviä haasteita lähes kaikissa ihmisen kognitiivista prosessia vastaavissa vaiheissa. Keskeisiä haasteita olivat:

- Ongelmatilanteen monimutkaisuuden takia tavoitetilan määrittely on vaikeaa
- Ongelmanratkaisun edellyttämien kaikkien mahdollisten tilojen määrittely on työlästä tai mahdotonta
- Tarvittavan tiedon mallintaminen on haastavaa
- Oleellinen tieto voi olla puutteellista, virheellistä tai vinoutunutta

Teoreettisen viitekehyksen pohjalta aineistossa esitetyt keskeiset ongelmat liittyvät erityisesti frame-ongelmaan ja tiedon hankinnan ongelmaan. Frame-ongelman osalta aineistosta nousi esiin, että tekoälyllä on haasteita sopeutua monimutkaisiin ja epämääräisiin ongelmanratkaisutilanteisiin. Nykyinen tekoäly kykenee pääasiassa ratkomaan ongelmia vain rajatuissa ja ennalta määritellyissä tilanteissa. Frame-ongelmaan liittyy myös haaste kokonaiskuvan hahmottamisesta; kaikkien näkökulmien huomioon ottaminen on haastavaa ja kaikkien mahdollisten tilojen mallintaminen on työlästä tai mahdotonta.

Tiedon hankinnan ongelmaan liittyy keskeisesti tarpeellisen ja oleellisen tiedon tunnistaminen ja saatavuus. Nykyiset tekoälyratkaisut tarvitsevat toimintansa pohjalle koulutusaineiston tai tietokannan, jota voidaan hyödyntää ongelmanratkaisutilanteissa. Tutkimusaineistossa korostui, että tietokantojen tuottaminen on työlästä ja tiedon muuttaminen ja luokittelu tekoälyn hyödynnettävissä olevaan muotoon on haastavaa. Lisäksi tiedon hyödynnettävyyteen liittyy luotettavuusongelmia: tieto voi olla virheellistä, vinoutunutta tai puutteellista.

Frame-ongelmaa tai tiedon hankinnan ongelmaa ei sellaisenaan mainittu tutkimusaineistossa, vaan olen itse tulkinnut aineistosta nousseet haasteet niiden kautta. Omasta mielestäni on yllättävää, että tutkimusaineistossa tekoälyn ongelmia ei käsitelty mainittujen tai muiden aiemmin tunnistettujen ongelmien näkökulmasta, vaan ne lueteltiin ainoastaan yleisinä huomioina. Kuvastaako tämä mahdollisesti tutkimusalan konsensuksen puutetta, tutkimusartikkeleiden teoreettisen taustoituksen heikkoutta vai kenties oman aineistoni puutteellisuutta. Itse koin tärkeänä kytkeä aineistosta esiin nousseet haasteet osaksi laajempaa teoreettista viitekehystä.

Teoriakatsauksen perusteella tekoälyn haasteisiin kuuluvat myös assosiaatiofunktion ongelma ja symbolisten merkitysten ongelma. Näihin ongelmiin liittyviä mainintoja en tutkimusaineistossa havainnut. Syitä niiden puuttumiselle voi vain arvailla: ovatko ne kenties vähemmän tutkittuja tai tunnettuja, onko niihin vaikeampi löytää toimivia ratkaisuja tai nähdäänkö ne vähemmän oleellisena

monimutkaisen ongelmanratkaisun kannalta. Toisaalta puuttuminen voi selittyä myös tutkimusaineiston rajallisuudella.

5.2 Ratkaisuja tekoälyn ongelmanratkaisukyvyyn kehittämiseksi

Toisena tutkimuskysymyksenä selvitin: *minkälaisia ratkaisuehdotuksia viime vuosina on esitetty sille, että tekoäly kykenisi ongelmanratkaisuun monimutkaisissa ja jäsenmättömissä tilanteissa?* Toisen tutkimuskysymyksen osalta aineistossa oli erilaisia ratkaisuehdotuksia tekoälyn ongelmanratkaisukyvyyn kehittämiseksi. Ratkaisuna esitettiin muun muassa siirto-oppimiseen, aktiiviseen oppimiseen ja kokemuspohjaiseen oppimiseen perustuvia malleja, mutta myös hakuavaruuden mallintamista ja mittaamista. Tekoälyn ongelmanratkaisuun liittyvien haasteiden osalta tutkimusaineistossa esitetyt ratkaisut vastaavat pääasiassa yhden tai kahden haasteen ratkaisemiseen, kuten kuviosta 9 voidaan havaita. Mikään yksittäinen ratkaisu ei vastaa kaikkiin tunnistettuihin haasteisiin. Kattavimman ratkaisuehdotuksen esittävät Hou ym. (2017) evolutiivisen siirtoavaruusoppimisen viitekehyksen muodossa, mutta sekään ei tulkintani mukaan ratkaise kaikkia tunnistettuja haasteita.

Keskeisimpänä eri ratkaisuja yhdistävänä tekijänä oli tulkintani mukaan ihmisen hyödyntäminen tekoälyn tukena ongelmanratkaisussa. Kahdeksasta artikkelista viidessä käsiteltiin jollain tasolla ihmisen ja tekoälyn välisen tiedonkäsittelyn yhdistämistä tai tekoälyn toiminnan tarkistamista ihmisen toimesta. Näkökulmat ihmisen ja tekoälyn yhteistyöhön liittyivät erityisesti ihmisen rooliin kouluttajana, aineiston validoijana tai ongelmatilanteen jäsentäjänä:

- *Ihminen toimii kouluttajana* - Järjestelmä kysyy uuden tilanteen edessä ihmiseltä neuvoa ongelman ratkaisemiseen (Moral, 2020). Kokemuksien kartuttaminen tapahtuu ihmisen ja koneen välisessä vuorovaikutuksessa (Mokhtari, 2019).
- *Ihminen jäsentää ongelmatilannetta* - Yhdistelmä tekoälyn laskenta-tehoa ja ihmisen kokonaisvaltaista tilannekäsitystä on tulevaisuuden tiedonkäsittelyn kehityssuunta (Li, 2020). Ihminen ja tekoäly työskentelevät yhdessä: ihminen huolehtii intuitiosta ja kehittää jatkuvasti haun suuntaa, kun taas tekoäly hyödyntää nopeuttiaan ja muistiaan tehdäkseen perusteellisia hakuja (Jones & Tanimoto, 2018).
- *Ihminen validoi aineiston* - Ihminen kehittää ja täydentää tekoälyä: toimii aineiston validoijana sekä tiedon lähteenä (Correira, 2019).

Ihmisen ja tekoälyn vahvuuksien yhdistäminen on varmasti perusteltu ja järkevä kehitysaskel tekoälyn ongelmanratkaisukyvyyn kehittämässä. Uskoisin, että tällaisia ratkaisuja on jo käytössä ja tämä on varmasti toimiva ja helppo kehityssuunta myös lähivuosien osalta. Toisaalta tekoälyn ongelmanratkaisukyvyyn

kehittäminen ihmisen roolia korostamalla tuntuu omasta mielestäni vajavaiselta kehitykseltä. Se saa minut pohtimaan

- a) onko jossain määrin luovuttu ajatuksesta, että ihminen pystyy mallintamaan älykkyyden ja luomaan sellaisen keinotekoisesti ja
- b) ovatko pelkästään tekoälyn omaan toimintakykyyn perustuvat kokonaisvaltaiset ratkaisumallit toistaiseksi liian vaikeita kehittää.

Jäin pohtimaan tutkimusartikkeleiden pohjalta, riittääkö yhdenkään niissä esitetyistä teorioista eteenpäin kehittäminen tuottamaan monimutkaiseen ongelmanratkaisuun kykenevää tekoälyä. Hou ym. (2017) tutkimus evolutiivisesta siirtovahvistusoppimisesta on perusteellinen ja ratkaisuehdotus vaikuttaa muita tutkimusaineiston ratkaisuja monipuolisemmalla. Erityisen kiinnostavaa on, että artikkelissa tarkastellaan sekä meemien hyödyntämistä, että evolutiivista oppimisprosessia. Pidän silti epätodennäköisenä, että yhdellä mallilla pystyttäisiin kattavasti ratkaisemaan kaikki nykyisen tekoälyn ongelmanratkaisun haasteet. Myös Marcus ja Davis (2019) sanovat, että mikään yksittäinen tekniikka ei pysty ratkaisemaan niitä asioita, joiden parissa tutkimuskenttä on ponnistellut vuosia. Tunnistamieni haasteiden ja niihin esitettyjen ratkaisujen perusteella ajattelen Marcusin ja Davisin tapaan, että tekoälyn kehittämisen kannalta toimivin lopputulos on luoda kokonaisuus, jossa yhdistyy useita erilaisia ratkaisuja. Yksikään nykyisistä tekoälysovelluksista ei sellaisenaan kykene ratkomaan monimutkaisia ja jäsentymättömiä ongelmia, mutta kenties tekoäly, joka osaisi valita eri tilanteisiin sopivan oppimisen, tapauskohtaisen päättelyn ja luonnollisen kielen käsittelyn tekniikan voisi kyetä jo nykyistä edistyneempään ongelmanratkaisuun. Toki onnistumiset tietyillä ongelma-alueilla (esim. ongelmatilanteen määrittely) voivat edistää myös muiden haasteiden ratkaisemista. Näkisin, että tulevaisuuden tekoälysovellus voisi esimerkiksi olla hierarkkinen kokonaisuus, jossa tekoäly arvioi ensin ongelmatilanteen ja sen jälkeen valitsee tilanteessa hyödynnettävän ongelmanratkaisutekniikan.

Tutkimusaineiston perusteella on mahdollista sanoa, mitkä ovat todellisuudessa tekoälyn ongelmanratkaisussa seuraavat kehitysaskeleet. Historiallista näkökulmasta tarkasteluna nykyhetken tekoälyinnostusta saattaa seurata jälleen yksi tekoälytalvi. Tämä saattaisi olla seurauksena esimerkiksi siitä, että yleisiin kielimalleihin kohdistuvia suuria odotuksia ei kyetä lunastamaan. Tätä kirjoittaessa esimerkiksi Chat GPT:n taustalla olevassa OpenAI:ssa kuohuu toimitusjohtajan vaihdosten merkeissä.

5.3 Tutkimustulosten merkitys

Tekoälyn kehitysnäkymien kartoittamiseksi koin perustelluksi selvittää tässä tutkimuksessa, mitä keskeisiä ongelmia tekoälyn ongelmanratkaisukykyyn osalta tällä hetkellä tunnustetaan ja minkälaisia ratkaisuehdotuksia niihin esitetään. Tutkimukseni toimii siten myös ihmisen ja tekoälyn älykkyyttä jäsentävänä katsauksena, sillä ongelmanratkaisu on yksi älykkyyden keskeisistä muodoista.

Nämä ovat keskeisiä näkökulmia muun muassa arvioitaessa, kuinka kaukaisia visioita yleinen tekoäly tai superäly käytännössä ovat.

Tiedostin ja tutkimukseni alkuvaiheessa, että tekoälyn ongelmanratkaisu on laaja aihe. Vaikka yritin miettiä tutkimusmenetelmän ja tutkimuskysymykset sopivasti aihetta rajaaviksi, oli kokonaisuus silti edelleen erittäin laaja. Tutkimusongelman laajuudesta ja toisaalta käytetyissä hakutermeissä tunnistetusta ongelmasta (ks. kappale 3.7.2) johtuen tätä tutkimusta ei voi pitää kattavana katsauksena tekoälyn ongelmanratkaisun haasteisiin ja mahdollisiin ratkaisuihin. Aiheen laajuudesta ja tutkimuksen rajallisuudesta huolimatta uskon pystyneeni taustoittamaan aihetta havainnollisesti ja käsittelemään tutkimusongelmaa koskevia teemoja jäsennellysti.

Vaikka kirjallisuuskatsauksen hakutermit osoittautuivat myöhemmässä tarkastelussa hieman ongelmallisiksi, kirjallisuuskatsaus on toteutettu systemaattisen kirjallisuuskatsauksen periaatteiden mukaisesti ja kuvattu huolellisesti. Systemaattinen kirjallisuuskatsaus on mielestäni ollut toimiva menetelmä selvittää vastauksia tutkimuskysymyksiini. Kirjallisuuskatsauksen kautta olen voinut lähestyä tekoälyn ongelmanratkaisua nykytilanteen haasteiden sekä tulevaisuuden mahdollisuuksien näkökulmasta. Kuten Snyder (2019) toteaa, kirjallisuuskatsauksella on tärkeä rooli kaikytyypisten tutkimusten pohjana. Ne voivat toimia pohjana tiedon kehittämiseksi, luoda suuntaviivoja uusille menettelytavoille ja käytännöille, tuottaa todistusaineistoa vaikutuksista ja hyvin toteutettuna myös luoda uusia ideoita ja suuntia tutkimuskentälle. (Snyder, 2019.)

Olen tutkimuksen tulokset kappaleessa esitellyt ratkaisuehdotukset suhteessa tunnistettuihin ongelmiin. Ratkaisut olisi voinut jaotella myös muilla tavoin, kuten menetelmien pohjalta tai ratkaisujen keskinäisten yhteneväisyyksien perusteella. Mielestäni ongelman ja ratkaisun yhdistäminen oli luonteva tapa kuvata tuloksia. Johtopäätökset -kappaleessa esittämäni synteesi (kuvio 9) on oma kontribuutioni tutkimusaiheeseen ja sitä voidaan hyödyntää esimerkiksi tulevien tutkimusten pohjana. Näkemykseni mukaan tekoälyä koskeva tutkimuskenttä hyötyisi siitä, että käsitteistä ja keskeisistä teorioista saataisiin luotua vakiintuneempia yhtenäisiä näkemyksiä. Toki voi olla utopistinen tavoite, että niin monisyisille ilmiöille ja käsitteille kuin älykkyys ja ongelmanratkaisu pystyttäisiin muotoilemaan täysin kattavia määritelmiä. Tutkijan työtä kuitenkin helpottaisi, jos alalle saataisiin luotua jonkinlaista vakiintunutta näkemystä. Koen, että olen onnistuneesti koonnut yhteen ja tiivistänyt osan alan viimeisintä tutkimustietoa. Toisaalta katsauksen suppeutta kuvastaa se, että yhdessäkään tutkimusaineiston artikkeleista ei käsitelty luonnollisen kielen käsittelymalleja, vaikka ne ovat olleet merkittävän huomion ja kiinnostuksen kohteena viime vuosien aikana.

5.4 Mahdollisia jatkotutkimuksia

Tutkimusaineiston perusteella vaikuttaa siltä, että kaiken kattavaa ratkaisua tekoälyn ongelmanratkaisukyvyyn kehittämiseksi ei vielä ole löytynyt. Lienee siis pitkä matka siihen, että tekoäly kykenee ihmisen kaltaiseen monimutkaiseen ongelmanratkaisuun jäsentymättömässä ympäristössä. Ihmisen kaltaiseen

älykkääseen toimintaan kykenevän yleisen tekoälyn kehittäminen edellyttää useiden haasteiden ratkaisemista sekä todennäköisesti useampia ratkaisuja yhdistävän kokonaisuuden kehittämistä.

Tämän tutkimuksen tulosten yleistettävyyden kartoittamiseksi olisi hyvä toteuttaa toinen kirjallisuuskatsaus, jossa hakutermit olisi määritelty huomioiden tässä tutkimuksessa havaitut haasteet. Kirjallisuuskatsauksen voisi myös toteuttaa tätä tutkimusta laajemmassa mittakaavassa, jolloin mahdollisuus saturaatioon olisi suurempi. Mikäli katsaukseen haluaisi vahvemmin kognitiotieteellisen näkökulman, seulontakriteeriksi voisi lisätä vaatimuksen, että ratkaisuehdotus on johdettu ihmisen älykkyyden pohjalta.

Tutkimusta tekoälyn ongelmanratkaisusta voisi myös jatkaa selvittämällä, miten aineistossa esitetyt ratkaisuehdotuksia on muualla tutkittu ja minkälaisia sovellusmahdollisuuksia niiden osalta on tunnistettu. Tarkempaan tarkasteluun voisi ottaa esimerkiksi 1–5 tässä tutkimuksessa tunnistetuista ratkaisuehdotuksista. Tällaisessa tutkimuksessa olisi hyvä muistaa, että yksittäisellä ratkaisuehdotuksella tuskin löydetään kaikki tekoälyn haasteet ratkaisevaa mallia.

Vaihtoehtoisesti tutkimuksen pohjalta voisi kartoittaa, mitä mieltä alan asiantuntijat ovat tutkimuksessa tunnistetuista ratkaisuehdotuksista ja minkälaisia mahdollisuuksia tai heikkouksia he tunnistavat niissä. Tällaisen jatkotutkimusten pohjalta olisi mahdollista syventää tutkimusongelmaa.

Mahdollista empiiristä tutkimusta aiheesta voisi toteuttaa esimerkiksi ihmisen ja tekoälyn ongelmanratkaisukykyjä vertailemalla. Tutkimusaineiston perusteella seuraavana kehityssakeleena tulee todennäköisesti olemaan ihmisen ja koneen yhteistyö, jossa ongelmanratkaisu tapahtuu molempien vahvuuksia hyödyntäen, joten ihmisen ja tekoälyn ominaisuuksia kartoittavalle tutkimukselle olisi tarvetta.

Kuten johdannossa totesin, tekoälyn ongelmanratkaisukyvyn ymmärtäminen on tärkeää, jotta järjestelmistä voidaan kehittää luotettavia, turvallisia, toimivia ja tuloksellisia. Tämä tutkimus on osaltaan osoittanut, ettei tutkimusalalla vallitse yhteistä käsitystä tekoälyn ongelmanratkaisukyvyn mahdollisuuksista ja rajoista. Aiheesta tarvitaan edelleen kattavaa ja monitieteellistä tutkimusta. Tässä tutkimuksessa tuotettua viitekehystä (kuvio 9) on mahdollista hyödyntää pohjana tutkimusongelman tarkastelussa jatkossa.

LÄHTEET

- Baum, E. (2007). A Working Hypothesis for General Intelligence. Teoksessa *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms: Proceedings of the AGI Workshop 2006*. IOS Press, Incorporated.
- Bhandari, P. (2022). Triangulation in Research | Guide, Types, Examples. Scribbr. <https://www.scribbr.com/methodology/triangulation/>
- Bian, N., Han, X., Sun, L., Lin, H., Lu, Y., & He, B. (2023). *ChatGPT is a Knowledgeable but Inexperienced Solver: An Investigation of Commonsense Problem in Large Language Models* (arXiv:2303.16421). arXiv. <https://doi.org/10.48550/arXiv.2303.16421>
- Breuker, J. (2013). A cognitive science perspective on knowledge acquisition. *International Journal of Human-Computer Studies*, 71(2), 177–183. <https://doi.org/10.1016/j.ijhcs.2012.10.006>
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). *Sparks of Artificial General Intelligence: Early experiments with GPT-4* (arXiv:2303.12712). arXiv. <https://doi.org/10.48550/arXiv.2303.12712>
- Buchanan, B. G. (2005). A (Very) Brief History of Artificial Intelligence. *AI Magazine*, 26(4), Article 4. <https://doi.org/10.1609/aimag.v26i4.1848>
- Buligi, I., Ciora, L. I., Petcusin, F., Badica, A., & Badica, C. (2020). Quantifying Blocks World State Space. *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, 1–7. <https://doi.org/10.1109/INISTA49547.2020.9194674>
- Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., Wang, Y., Ye, W., Zhang, Y., Chang, Y., Yu, P. S., Yang, Q., & Xie, X. (2023). *A Survey on Evaluation of Large Language Models* (arXiv:2307.03109). arXiv. <http://arxiv.org/abs/2307.03109>
- Condell, J., Wade, J., Galway, L., McBride, M., Gormley, P., Brennan, J., & Somasundram, T. (2010). Problem solving techniques in cognitive science. *Artificial Intelligence Review*, 34(3), 221–234. <https://doi.org/10.1007/s10462-010-9171-0>
- Confalonieri, R., Coba, L., Wagner, B., & Besold, T. R. (2021). A historical perspective of explainable Artificial Intelligence. *WIREs Data Mining and Knowledge Discovery*, 11(1), e1391. <https://doi.org/10.1002/widm.1391>
- Conway, A. R. A., & Kovacs, K. (2015). New and emerging models of human intelligence. *WIREs Cognitive Science*, 6(5), 419–426. <https://doi.org/10.1002/wcs.1356>
- Correia, A., Paredes, H., Schneider, D., Jameel, S., & Fonseca, B. (2019). Towards Hybrid Crowd-AI Centered Systems: Developing an Integrated Framework from an Empirical Perspective. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 4013–4018. <https://doi.org/10.1109/SMC.2019.8914075>
- Dennett, D. (1984). *Cognitive Wheels: The Frame Problem of AI*.
- Dietrich, E., & Fields, C. (2020). Equivalence of the Frame and Halting Problems. *Algorithms*, 13(7), 175. <https://doi.org/10.3390/a13070175>
- Edinger, A., & Goldstone, R. (2022). Getting Situated: Comparative Analysis of Language Models With Experimental Categorization Tasks. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 44(44). <https://escholarship.org/uc/item/2fd1x2gz>
- Elo, S., Kajula, O., Tohmola, A., & Kääriäinen, M. (2022). Laadullisen

- sisällönanalyysin vaiheet ja eteneminen. *Hoitotiede*, 34(4), Article 4.
- Fink, A. (2010). *Conducting Research Literature Reviews: From the Internet to Paper*. SAGE Publications.
- Floridi, L. (2020). AI and Its New Winter: From Myths to Realities. *Philosophy & Technology*, 33(1), 1–3. <https://doi.org/10.1007/s13347-020-00396-6>
- Ford, M. (2018). *Architects of Intelligence: The Truth About AI From the People Building It*. Packt publishing.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1–38. [https://doi.org/10.1016/0010-0285\(83\)90002-6](https://doi.org/10.1016/0010-0285(83)90002-6)
- Goldstein, E. B. (2015). *Cognitive psychology: Connecting mind, research, and everyday experience* (4th edition, [student edition]). Cengage Learning.
- Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1), 13–23. [https://doi.org/10.1016/S0160-2896\(97\)90011-8](https://doi.org/10.1016/S0160-2896(97)90011-8)
- Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). *When Will AI Exceed Human Performance? Evidence from AI Experts* (arXiv:1705.08807). arXiv. <https://doi.org/10.48550/arXiv.1705.08807>
- Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), 5–14. <https://doi.org/10.1177/0008125619864925>
- Harnad, S. (1990). The Symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1–3), 335–346.
- Hayashi, P., Abib, G., & Hoppen, N. (2019). The Qualitative Report Validity in Qualitative Research: A Processual Approach. *Qualitative Report*, 24, 98–112. <https://doi.org/10.46743/2160-3715/2019.3443>
- Hiebl, M. R. W. (2023). Sample Selection in Systematic Literature Reviews of Management Research. *Organizational Research Methods*, 26(2), 229–261. <https://doi.org/10.1177/1094428120986851>
- Hirsjärvi, S., Remes, P., Sajavaara, P., & Sinivuori, E. (1997). *Tutki ja kirjoita* (1.-2. p). Tammi.
- Hoffmann, C. H. (2022). Is AI intelligent? An assessment of artificial intelligence, 70 years after Turing. *Technology in Society*, 68, 101893. <https://doi.org/10.1016/j.techsoc.2022.101893>
- Holth, P. (2008). What is a problem? Theoretical conceptions and methodological approaches to the study of problem solving. *European Journal of Behavior Analysis*, 9(2), 157–172. <https://doi.org/10.1080/15021149.2008.11434302>
- Hossain, M. A., Md. Noor, R., Yau, K.-L., Razalli, S., Zraba, M., & Ahmedy, I. (2020). Comprehensive Survey of Machine Learning Approaches in Cognitive Radio-Based Vehicular Ad Hoc Networks. *IEEE Access*, 8, 78054–78108. <https://doi.org/10.1109/ACCESS.2020.2989870>
- Hou, Y., Ong, Y.-S., Feng, L., & Zurada, J. M. (2017). An Evolutionary Transfer Reinforcement Learning Framework for Multiagent Systems. *IEEE Transactions on Evolutionary Computation*, 21(4), 601–615. <https://doi.org/10.1109/TEVC.2017.2664665>
- Hwang, K. S., Park, K. S., Lee, S. H., Kim, K. I., & Lee, K. M. (2018). Autonomous Machine Learning Modeling using a Task Ontology. *2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS)*, 244–248. <https://doi.org/10.1109/SCIS-ISIS.2018.00051>

- Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). Quo vadis artificial intelligence? *Discover Artificial Intelligence*, 2(1), 4. <https://doi.org/10.1007/s44163-022-00022-8>
- Johnson, M., Albizri, A., Harfouche, A., & Fosso-Wamba, S. (2022). Integrating human knowledge into artificial intelligence for complex and ill-structured problems: Informed artificial intelligence. *International Journal of Information Management*, 64, 102479. <https://doi.org/10.1016/j.ijinfomgt.2022.102479>
- Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, 48(4), 63–85. <https://doi.org/10.1007/BF02300500>
- Jones, B. T., & Tanimoto, S. L. (2018). Searching Over Search Trees for Human-AI Collaboration in Exploratory Problem Solving: A Case Study in Algebra. *2018 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 33–37. <https://doi.org/10.1109/VLHCC.2018.8506580>
- Kolak, D., Hirstein, W., Mandik, P., & Waskan, J. (2006). *Cognitive science: An introduction to mind and brain*. Routledge.
- Korteling, J. E. (Hans)., van de Boer-Visschedijk, G. C., Blankendaal, R. A. M., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human- versus Artificial Intelligence. *Frontiers in Artificial Intelligence*, 4. <https://www.frontiersin.org/articles/10.3389/frai.2021.622364>
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40, e253. <https://doi.org/10.1017/S0140525X16001837>
- Langley, P., & Rogers, S. (2005). An Extended Theory of Human Problem Solving. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 27(27).
- Legg, S., & Hutter, M. (2007). A Collection of Definitions of Intelligence. Teoksessa *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms: Proceedings of the AGI Workshop 2006*. IOS Press, Incorporated.
- Lensu, A. (2002). *Computationally intelligent methods for qualitative data analysis*.
- Li, J., Cheng, H., Guo, H., & Qiu, S. (2018). Survey on Artificial Intelligence for Vehicles. *Automotive Innovation*, 1(1), 2–14. <https://doi.org/10.1007/s42154-018-0009-9>
- Li, W., Suh, C. S., Xu, X., & Song, Z. (2022). Extenics enhanced axiomatic design procedure for AI applications. *AI EDAM*, 36, e23. <https://doi.org/10.1017/S0890060422000075>
- Li, X., Wu, W., Wang, H., & Pi, H. (2020). How to solve Ill-defined problems Intelligently in the Big Data Environment. *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 922–926. <https://doi.org/10.1109/WIAT50758.2020.00140>
- Liu, H., Ning, R., Teng, Z., Liu, J., Zhou, Q., & Zhang, Y. (2023). *Evaluating the Logical Reasoning Ability of ChatGPT and GPT-4* (arXiv:2304.03439). arXiv. <https://doi.org/10.48550/arXiv.2304.03439>
- Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2018). Brain Intelligence: Go beyond Artificial Intelligence. *Mobile Networks and Applications*, 23(2), 368–375. <https://doi.org/10.1007/s11036-017-0932-8>
- Lucci, S., Musa, S. M., & Kopec, D. (2015). *Artificial Intelligence in the 21st Century*. Mercury Learning & Information. <http://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=7025176>
- Luger, G. F. (2005). *Artificial intelligence: Structures and strategies for complex*

- problem solving* (5th ed). Addison-Wesley.
- Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Pantheon Books.
- Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). *AI Index Report 2023*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA,.
- McCarthy, J., Minsky, M. L., Rochester, N., Corporation, I. B. M., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Magazine*, 27(4).
- McKinnon, J. (1988). Reliability and Validity in Field Research: Some Strategies and Tactics. *Accounting, Auditing & Accountability Journal*, 1(1), 34–54.
<https://doi.org/10.1108/EUM0000000004619>
- Miao, Q., Zheng, W., Lv, Y., Huang, M., Ding, W., & Wang, F.-Y. (2023). DAO to HANOI via DeSci: AI Paradigm Shifts from AlphaGo to ChatGPT. *IEEE/CAA Journal of Automatica Sinica*, 10(4), 877–897.
<https://doi.org/10.1109/JAS.2023.123561>
- Mokhtari, V., Manevich, R., Lopes, L. S., & Pinho, A. J. (2019). Learning the Scope of Applicability for Task Planning Knowledge in Experience-Based Planning Domains. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 3973–3979. <https://doi.org/10.1109/IROS40897.2019.8968013>
- Moral, R. C. B., Paulus, G. B., Assuncao, J. V. C., & Silva, L. A. L. (2020). Investigating Case Learning Techniques for Agents to Play the Card Game of Truco. *2020 19th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, 107–116. <https://doi.org/10.1109/SBGames51465.2020.00024>
- Mukhamediev, R. I., Popova, Y., Kuchin, Y., Zaitseva, E., Kalimoldayev, A., Symagulov, A., Levashenko, V., Abdoldina, F., Gopejenko, V., Yakunin, K., Mukhamedijeva, E., & Yelis, M. (2022). Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges. *Mathematics*, 10(15), Article 15.
<https://doi.org/10.3390/math10152552>
- Norris, N. (1997). Error, bias and validity in qualitative research. *Educational Action Research*, 5(1), 172–176. <https://doi.org/10.1080/09650799700200020>
- Norvig, P., & Russell, S. (Toim.). (2010). *Artificial Intelligence: A Modern Approach* (3rd. ed.). Prentice hall. <http://repo.darmajaya.ac.id/3800/1/Artificial%20Intelligence%20A%20Modern%20Approach%20%283rd%20Edition%29.pdf%20%28%20PDFDrive%20%29.pdf>
- Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *Social Science Research Network*.
<https://doi.org/10.2139/ssrn.1954824>
- Oswald, M. E., & Grosjean, S. (2012). Confirmation bias. *Teoksessa Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*. Psychology Press.
- Reeves, L., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, 115(3), 381–400.
<https://doi.org/10.1037/0033-2909.115.3.381>
- Russell, S. (2020). *Human Compatible: AI and the Problem of Control*. Penguin books.
- Shushan, C., & Xiaodi, X. (2017). *Artificial Intelligence and Human Intelligence*.
- Smith, M. U. (1991). *Toward a unified theory of problem solving: Views from the*

- content domains*. Lawrence Erlbaum, Associates, Inc.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Storås, N. (2023, huhtikuuta 19). *Tekoäly | Chat GPT:n takana on yhtiö, joka aikoo mullistaa nykyisen talousjärjestelmän*. Helsingin Sanomat. <https://www.hs.fi/viisio/art-2000009520008.html>
- Thagard, P. (2005). *Mind: Introduction to cognitive science* (2nd ed). MIT Press.
- Tieteen termipankki. (2023). *Epädeterministinen*. <https://tieteentermipankki.fi/wiki/Nimitys:ep%C3%A4deterministinen>
- Uddin, M. N. (2019). Cognitive science and artificial intelligence: Simulating the human mind and its complexity. *Cognitive Computation and Systems*, 1(4), 113–116. <https://doi.org/10.1049/ccs.2019.0022>
- Wang, Y., & Chiew, V. (2010). On the cognitive process of human problem solving. *Cognitive Systems Research*, 11(1), 81–92. <https://doi.org/10.1016/j.cogsys.2008.08.003>
- Watson, I. (1999). Case-based reasoning is a methodology not a technology. *Knowledge-Based Systems*, 12(5), 303–308. [https://doi.org/10.1016/S0950-7051\(99\)00020-9](https://doi.org/10.1016/S0950-7051(99)00020-9)

LIITE 1 ESIMERKKEJÄ HAKUTERMIEN TESTAUKSESTA

1)

Etsi artikkeleita kaikilla sanoilla	cognitive science artificial intelligence problem solving future
sisältäen tarkan ilmauksen	complex problem solving
ainakin yksi sanoista	artificial intelligence
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä	2017-2023

5130 hakutulosta

2)

Etsi artikkeleita kaikilla sanoilla	ai complex and ill-structured problems
sisältäen tarkan ilmauksen	problem solving
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä	2017-2023

3770 hakutulosta

3)

Etsi artikkeleita kaikilla sanoilla	cognitive science complex problem solving
sisältäen tarkan ilmauksen	artificial intelligence problem solving
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa

	<input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä	2017-2023

61 hakutulosta

4)

Etsi artikkeleita kaikilla sanoilla	cognitive science complex problem solving
sisältäen tarkan ilmauksen	artificial intelligence complex problem solving
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä	2017-2023

0 hakutulosta

5)

Etsi artikkeleita kaikilla sanoilla	cognitive science solutions for ai problem solving
sisältäen tarkan ilmauksen	artificial intelligence problem solving
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä	2017-2023

48 hakutulosta

6)

Etsi artikkeleita kaikilla sanoilla	cognitive science solutions for ai complex problem solving
sisältäen tarkan ilmauksen	artificial intelligence problem solving
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä	2017-2023

45 hakutulosta

7)

Etsi artikkeleita kaikilla sanoilla	cognitive science solutions for ai complex and ill-defined problem solving
sisältäen tarkan ilmauksen	artificial intelligence problem solving
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä:	2017-2023

4 hakutulosta

8)

Etsi artikkeleita kaikilla sanoilla	ai abilities to solve problems in ill-defined and complex situations
sisältäen tarkan ilmauksen	
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa

Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä:	2017-2023

17 200 hakutulosta

9)

Etsi artikkeleita kaikilla sanoilla	cognitive science ai abilities to solve problems in complex situations
sisältäen tarkan ilmauksen	
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä:	2017-2023

8980 hakutulosta

10)

Etsi artikkeleita kaikilla sanoilla	cognitive science ai abilities to solve problems in complex situations
sisältäen tarkan ilmauksen	artificial intelligence problem solving
ainakin yksi sanoista	
ilman sanoja	
jossa omat sanani esiintyvät	<input checked="" type="radio"/> artikkelin missä tahansa kohdassa <input type="radio"/> artikkelin otsikossa
Näytä artikkelit, jotka on kirjoittanut	-
Näytä artikkelit, jotka on julkaistu kohteissa	-
Näytä artikkelit, jotka on päivätty seuraavalla aikavälillä:	2017-2023

40 hakutulosta

LIITE 2 HAKUTULOSTEN TIETEELLISEN LAADUN ARVIOINTI

Tutkimusten tieteellisen laadun kartoittamiseksi pyrin tiivistelmän ja johtopäätösten pohjalta arvioimaan tutkimusten johdonmukaisuutta ja laatua. Lisäksi selvitin julkaisujen luokittelut julkaisufoorumissa. Luokittelutiedot löytyivät 13 tarkastelussa olleesta artikkelista kuuden osalta, joista yhden luokitus oli 3 eli korkein taso, neljän luokitus oli 1 eli perustaso (tosin yksin julkaisija on vuodesta 2023 ollut osa yritystä, jonka luokitus on 3) ja yhden luokitus 0 eli se ei ole täyttänyt tasolle 1 vaadittavia kriteereitä. Muiden julkaisujen osalta julkaisufoorumissa ei ollut tietoja. Yhden artikkelin hylkäsin arvioituani itse, ettei se vastaa tieteelliseltä artikkelilta edellytettävää laatua. Muiden artikkeleiden osalta yritin selvittää julkaisujen luotettavuutta myös avoimella haulla, minkä kautta sain tietää seuraavat julkaisijatiedot:

- ResearchGate on eurooppalainen kaupallisesti toimiva sosiaalinen verkkoalusta, jossa tutkijat voivat jakaa julkaisuja sekä esittää ja vastata kysymyksiin.
- Al-Mukhtar Journal of Sciences on Omar Al-Mukhtar Universityn vertaisarvioitu ja avoimesti saatavilla oleva verkkojulkaisu
- INISTA on konferenssi, jota on pidetty vuodesta 2005. Se keskittyy älykkäisiin järjestelmiin ja niihin liittyviin sovelluksiin teorian ja käytännön näkökulmasta
- SBGames on latinalaisen Amerikan suurin ja tärkein pelien ja digitaalisen viihteen tapahtuma, johon osallistuu vuosittain 1000 henkeä
- IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology on vuodesta 2001 järjestetty konferenssi, joka kokoaa yhteen eri alojen tutkijoita ja toimijoita. Konferenssissa esitellään älyteknologiaa koskevia alkuperäisiä tutkimustuloksia sekä jaetaan ajankohtaista tietoa viimeaikaisesta kehitystyöstä.
- International Conference on Soft Computing and Intelligent Systems and International Symposium on Advanced Intelligent Systems (SCIS&ISIS) on Aasiassa pidettävä kansainvälinen ”pehmeän laskennan” konferenssi. Konferenssin esitykset vertaisarvioidaan etukäteen alkuperäisyyden ja laadun osalta.

Edellä mainituista päätin rajata pois ResearchGatessa sekä Al-Mukhtar Journalissa julkaistut artikkelit, koska ensimmäisessä on mahdollista julkaista vertaisarvioimattomiakin artikkeleita ja jälkimmäinen on yksittäisen yliopiston julkaisu.

LIITE 3 ENSIMMÄINEN HAKUTULOS

Nro	Tyyppi	Otsikko	Tekijät	Julkaisu- vuosi	Julkaisija
1	Artikkeli	Introduction to Artificial Intelligence and Problem solving	Sharma, K. & Bhargav, J.	2022	International Journal of Scientific Research & Engineering Trends
2	Artikkeli toimitetussa teoksessa	Teaching of general psychology: problem solving	Gibson, D., Ifenthaler, D., Greiff, S.	2023	Springer, Cham
3	Artikkeli toimitetussa teoksessa	Modern AI and how we got here	Luger, G.F.	2021	Springer, Cham
4	E-kirja	Artificial intelligence in the 21st century	Lucci, S., Musa, S. & Kopec, D.	2022	Mercury Learning & Information
5	Artikkeli toimitetussa teoksessa	Mathematical Analysis: How Would Ai Tackle Math Olympiad Problems?	Yang, S.	2021	International Journal of High School Research
6	Artikkeli	Computer Problem-Solving by Applying Depth-First Search	San, T. & Bo, AA.	ei tietoa	Computer University, Magway
7	Kirja	Cognitive science: an introduction to the study of mind	Friedenberg, J., Silverman, G. & Spivey, M.	2021	SAGE Publications
8	Konferenssi-julkaisu	A 3D Learning Game for Representing Artificial Intelligence Problems	Derwich, S., Essalmi, F.	2017	Springer, Singapore
9	Konferenssi-julkaisu	Applied Aspects of the Integrated Problem Solving System with Natural Language Interface	Kurbatov, S., Fominykh, I. & Vorobev, a.	2020	V International Conference on Information Technologies in Engineering Education
10	Gradu	Artificial Intelligence: Algorithms, Operational Environments and Hyperbole	Griesmyer, D.	2018	U.S Army Command and General Staff College
11	Artikkeli toimitetussa teoksessa	Toward an Active, Pragmatic, Model-Revising Realism	Luger, G.F.	2021	Springer, Cham
12	Konferenssi-julkaisu	AI Models in Oceanographic	Aiken, J. & Corchado Rodriguez, J	2022	ResearchGate
13	Artikkeli toimitetussa teoksessa	Association and Connectionist Approaches to AI	Luger, G.F.	2021	Springer, Cham
14	Artikkeli	Comparative Analysis of Search Algorithms in AI	Shafique, A., Sohail, F. & Islam Qadri, A.	2022	ResearchGate
15	Artikkeli toimitetussa teoksessa	User involvement in collaborative decision-making dialog	Nothdurft, F., Bercher, P., Behnke, G., Minker, W.	2017	Springer, Singapore
16	Artikkeli	The Choice of Exhaustive or Heuristic search to Solve AI Problems	Gatsh, D.	2019	Al-Mukhtar Journal of Sciences
17	Artikkeli	A computational theory of evaluation in creative design	Wiltgen, B. & Goel, A.	2019	IBM Journal of Research and Development
18	Artikkeli toimitetussa teoksessa	Artificial intelligence and Autonomous Car	Ariturk, M., Yavuz, S. & Allahviranloo, T.	2020	John Wiley & Sons, Incorporated

19	Artikkeli	Software tools for learning artificial intelligence algorithms	Stamenković, S., Jovanović, N., Vasović, B. et al.	2023	Artificial Intelligence Review
20	Konferenssi-julkaisu	Addressing challenges to problem complexity: Effectiveness of AI assistance during the design process	Song, B, Soria Zurita, NF, Nolte, H, Singh, H, Cagan, J, & McComb, C.	2021	Proceedings of the ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference
21	Artikkeli	Design and evaluation of a deep learning recommendation based augmented reality system for teaching programming and computational thinking	Lin, P. -H. & Chen, S. -Y.	2020	IEEE Access
22	Artikkeli	Survey on artificial intelligence for vehicles	Li, J., Cheng, H., Guo, H. & Qiu, S.	2018	Automotive Innovation
23	Konferenssi-julkaisu	Exploration and Practice of Intelligent Educational Accomplishment of Computer Majors in Vocational Colleges in the Era of Artificial Intelligence	Jinghua, S.	2021	International Conference on Internet, Education and Information Technology (IEIT)
24	Artikkeli	Towards Building an AI Curriculum for High School Students	Nisheva-Pavlova, M.	2022	CEUR Workshop Proceedings
25	Artikkeli toimitussa teoksessa	Explanatory AI for pertinent communication in autonomous systems	Pol, M., Desalles, JL., Diaconescu, A.	2020	Springer, Cham
26	Konferenssi-julkaisu	Quantifying Blocks World State Space	I. Buligiu, L. I. Ciora, F. Petcuşin, A. Bădică and C. Bădică,	2020	International Conference on INnovations in Intelligent Systems and Applications (INISTA)
27	Konferenssi-julkaisu	Investigating case learning techniques for agents to play the card game of Truco	Moral, R., Paulus, G., Assunção J. & Silva, L.	2020	19th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)
28	Artikkeli	Artificial Intelligence and User Experience in reciprocity: Contributions and state of the art	Virvou, M.	2023	Intelligent Decision Technologies
29	Konferenssi-julkaisu	TVDP: Translational visual data platform for smart cities	Kim, S., Al-farrarjeh, A., Constantinou, G. & Shahabi, C.	2019	IEEE 35th International Conference on Data Engineering Workshops (ICDEW)
30	Artikkeli	Perspectives on becoming an applied machine learning scientist	Rasiwasia, N.	2019	Computer
31	Artikkeli toimitussa teoksessa	Conceptual Framework for Digital Transformation in Higher Education an Emerging Trend	Kaur, K. & Chhibber, P.	2022	Chapman and Hall/CRC
32	Artikkeli toimitussa teoksessa	Learning problem-solving strategies in virtual worlds that encourage people to respect human rights	Bennett, K.	2019	IGI Global
33	Artikkeli	A framework for ethical artificial intelligence-from social theories to cybernetics-based implementation	Kushal, A.	2021	Inderscience Enterprises

34	Artikkeli toimitetussa teoksessa	PyTorch	Imambi, S., Prakash, K.B., Kanagachidambaresan, G.R.	2021	Springer, Cham
35	Gradu	On some applications of rough set theory in the study of knowledge discovery in datavases	Raghavan, R.	2021	ResearchGate
36	Artikkeli	An evolutionary transfer reinforcement learning framework for multiagent systems	Hou, Y., Ong, Y. -S., Feng, L. & Zurada, J.	2017	IEEE Transactions on Evolutionary Computation,
37	Artikkeli	Machine learning as a reusable microservice	Pahl, M. -O. & Loipfinger, M.	2018	IEEE/IFIP Network Operations and Management Symposium
38	Väitöskirja	Data-driven and machine learning based design creativity	Chen, L.	2019	Imperial College London
39	Konferenssi-julkaisu	Deepcloud. The application of a data-driven, generative model in design	Bidgoli, A. & Veloso, P.	2018	ACADIA 2018: Recalibration. On imprecision and infidelity. Proceedings of the 38th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)
40	Artikkeli	Fake News Detection using Semi-Supervised Graph Convolutional Network	Meel, P. & Vishwakarma, D.	2021	arXiv
41	Artikkeli	Advancing Smart Home Awareness-A coceptual Computational Modelling Farmework for the Execution of Daily Activities of People with Alzheimer's	Liappas, N.; Teriús-Padrón, J.G.; Garcia-Betances, R.I.; Cabrera-Umpiérrez, M.F.	2022	Sensors
42	Artikkeli toimitetussa teoksessa	1 The Significant Highlights of PyTorch	Prakash, K. & Kanagachidambaresan, G.	2021	Springer, Cham
43	Laskentatou-lukko	International Bar Association President's Task Force on the Future of Legal Services	Esteban, M. & Klotz, J.	2017	?
44	Väitöskirja	An End-To-End Cloud-Based Solution for Optimal Attention Network Topology in Real-Time Applications	Srraf, S.	2022	ProQuest
45	Väitöskirja	Optimal seismic Retrofitting of existing RC frames through soft-computing approaches	Falcone, R.	2018	Universita degli studi di Salerno
46	Gradu	A Swedish Natural Language Processing Pipeline For Building Knowledge Graphs	Gonzalez, A.	2018	Digitala Vetenskapliga Arkivet
47	Artikkeli toimitetussa teoksessa	Judgement, Decision-Making, and Intuition in the Pilot Selection Process	Hubbard, T. & Wolfe, P.	2019	CRC Press
48	Väitöskirja	A smart thermostat for demand-side management with real-time electricity prices	Shann, M.	2019	The Faculty of Business, Economics and Informatics of the University of Zurich

LIITE 4 TOINEN HAKUTULOS

Nro	Tyyppi	Otsikko	Tekijät	Julkaisu- vuosi	Julkaisija
1	Kirja	Fundamentals of Artificial Intelligence: Problem Solving and Automated Reasoning, 1st Edition	Kubat, M.	2023	McGraw Hill
2	Konferenssijulkaisu	How to Solve ill-defined problems Intelligently in the Big Data Environment	Li, X., Wu, W., Wang, H. & PL, H.	2020	IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)
3	Konferenssijulkaisu	Searching Over Search Trees for Human-AI Collaboration in Exploratory Problem Solving: A Case Study in Algebra	Jones, B. & Tanimoto, S.	2018	IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)
4	Konferenssijulkaisu	Decision-Making in Groundwater Management: Where Artificial Intelligence Can Really Lead Geoscientists?	Manziona, R.L. & Matulovic, M.	2021	Springer, Cham
5	Tiivistelmä tms	Keynote Speech V: Artificial Intelligence Computing Paradigms in Smart Healthcare Informatics	Salem, A. -B. M.	2019	9th International Conference on Computer Theory and Applications (ICCTA)
6	Uutiskirje	Association for automated reasoning		1997	
7	Artikkeli	The role of Artificial Intelligence in Online Dispute Resolution: A brief and critical overview	Alessa, H.	2022	Information & Communication Technology Law
8	Konferenssijulkaisu	Problem solving using social networks in Cultural Algorithms with auctions	Reynolds, R. & Kinnaird-Heether, L.	2017	IEEE International Conference on Humanized Computing and Communication (HCC)
9	Konferenssijulkaisu	Interdisciplinary Project Based Learning Approach for Machine Learning and Internet of Things	Khan, M., Ibrahim, M., Wu, N. & Patil, R.	2020	IEEE Integrated STEM Education Conference (ISEC)
10	Artikkeli	A Systematic Literature Review of Multi-agent Pathfinding for Maze Research	Tjiharjadi, S., Razali, S. & Sulaiman, H. A.	2022	Journal of Advances in Information Technology
11	Artikkeli toimitetussa teoksessa	AI: Methods and Techniques. Knowledge-Based Systems	Casas, D.L., Sierra, J.P.	2022	Palgrave Macmillan, Cham
12	Konferenssijulkaisu	Autonomous machine learning modeling using a task ontology	Hwang, K. S., Park, K. S., Lee, S. H., Kim, K. I. & Lee, K. M.	2018	Joint 10th International Conference on Soft Computing and Intelligent Systems and 19th International Symposium on Advanced Intelligent Systems
13	Väitöskirja	Development of an Integrated Artificial Intelligence Education Program Embedding Ethics	Bae, J.	2023	JEJU NATIONAL UNIVERSITY
14	Konferenssijulkaisu	Towards hybrid crowd-AI centered systems: Developing an integrated framework from an empirical perspective	Correia, A., Paredes, H., Schneider, D., Jameel, S. & Fonseca, B.	2019	IEEE International Conference on Systems, Man and Cybernetics (SMC)
15	Konferenssijulkaisu	A schematic review on application of deep learning and computer vision	Pitale, R., Kale, H., Kshirsagar, S. & Rajput, H.	2021	Asian Conference on Innovation in Technology (ASIANCON)
16	Konferenssijulkaisu	Intelligent Spatial Decision Support System	Yusianto, R. Marimin, Suprihatin &	2020	International Conference on Computer Science and Its Application in Agriculture (ICOSICA)

		Concept in the Potato Agro-Industry Supply Chain	Hardjomid-jojo, H.		
17	Artikkeli	Novel domain expansion methods to improve the computational efficiency of the Chemical Master Equation solution for large biological networks	Kosarwal, R., Kulasiri, D. & Samarasinghe, S.	2020	BMC Bioinformatics
18	Artikkeli	Alpha-beta pruning-a streamline approach for perceptive game palying	Vinay Kumar, B. C., Jashank, N., Manjari, S. N., Nagalakshmi, R. & Sanath, H. N.	2020	International Research Journal of Modernization in Engineering Technology and Science
19	Artikkeli	Intelligent sales volume forecasting using google searc angine data	Yuan, FC., Lee, CH.	2020	Soft Computing
20	Artikkeli	A parallel multi-objective cooperative co-evolutionary algorithm with changing variables	Xu, B., Zhang, Y., Gong, D.-W. & Wang, L.	2017	Gecco'17
21	Artikkeli	Automatic adaptation of basal insulin using sensor-augmented pump therapy	Herrero, P., Bondia, J., Giménez, M., Oliver, N. & Georgiou, P.	2018	Journal of Diabetes Science and Technology
22	Konferenssijulkaisu	Learning the scope of applicability for task planning knowledge in experience-based planning domains	Mokhtari, V., Manevich, R., Lopes, L. S. & Pinho, A. J.	2019	IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),
23	Konferenssijulkaisu	An approach to robot task learning and planning with loops	Mokhtari, V., Lopes, L. S. & Pinho, A. J.	2017	IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
24	Artikkeli	Artificial intelligence in Logistics. Applications and algorithms	Kota, L.	2018	Advanced Logistic Systems
25	Artikkeli toimittetussa teoksessa	Overview of Robots	Duan, F., Li, W., Tan, Y.	2023	Springer, Singapore.
26	Artikkeli toimittetussa teoksessa	Visualizing Markov Process Through Graphs and Trees	Kulasiri, D. & Kosarwal, R.	2021	Springer, Singapore
27	Artikkeli	Emerging Technological Innovations in the Legal Profession and Its Impact on the Regulation of Market Competition: Kenyan Perspective	Kigwiru, V. K.	2019	SSRN
28	Artikkeli	Optimal multi-way number partitioning	Schreiber, E., Korf, R. & Moffitt, M.	2018	Journal of ACM
29	Väitöskirja	Optimal sequential decision-making under uncertainty	Peron, M.	2018	Queensland University of Technology
30	Kirja	International tourism futures: The drivers and impacts of change	Lade, C., Strickland, P., Frew, E., Willard, P., Nagpal, S., Osorio, S. & Vitartas, P.	2020	Goodfellow publishers
31	Väitöskirja	A game theory approach for the collaborative planning of production and transportation activities in the supply chain	Wang, Y.	2018	Université de Bordeaux
32	Opetussuunnitelma	Research Methodology			

33	Väitöskirja	Aplicação das tecnologias de inteligência artificial em operações militares	Bettencourt, J.	2022	Institutio Universitadio militar
34	Artikkeli	Inteligencia artificial en el Derecho Internacional, Naciones Unidas y Unión Europea	Drnas de Clément, Z. .	2022	Universidad de Jaen